Project 1

SQL Queries for getting the data are:

```
Select * from city_data
where  city='Philadelphia';

Select * from city_list
where city='Philadelphia';

Select * from global_data;
```

Steps taken

I used the query above to get the data which I exported to CSV files. I looked at the data for city list and found the nearest city to me is Philadelphia.

I opened the file for the global data and for each year, I did VLOOKUP to get the corresponding Philadelphia temperature. Thus, each row of data now contains the year, global temperature and Philadelphia temperature. Then I saved the file.

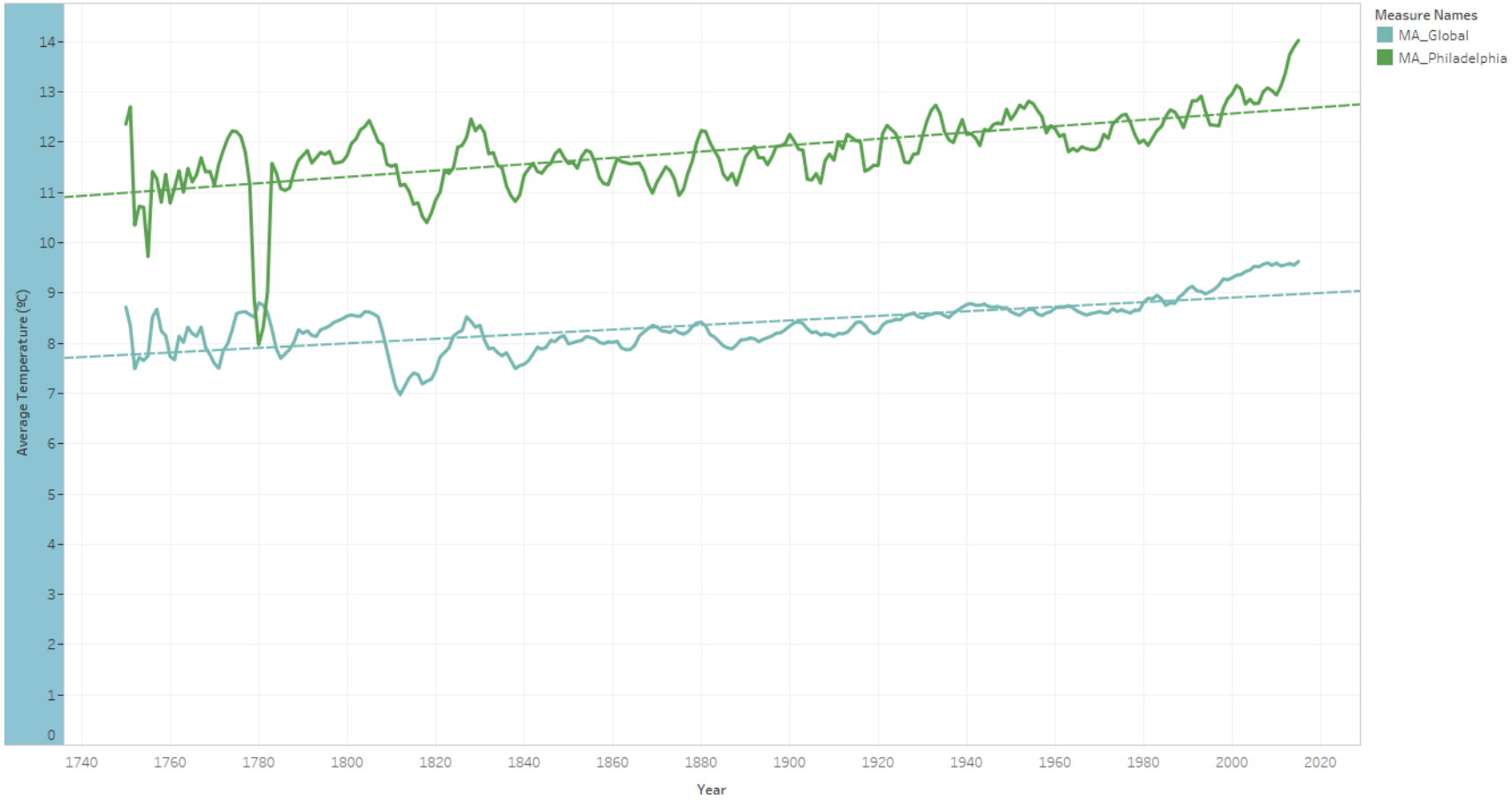
I chose to use Tableau to analyze the data. I imported the updated global data (including Philadelphia data) into Tableau. To smoothen the data, I decided to calculate the 3-year moving average for both the global temperature and Philadelphia temperature.

I calculated the moving average for the global temperature using the formula WINDOW_AVG(SUM([Global AVG Temp]), -3, 0) and Philadelphia moving average using the formula WINDOW_AVG(SUM([Philadelphia AVG Temp]), -3, 0). I named the former MA_Global and the latter MA_Philadelphia.

To plot a line chart,

- I placed the Year on the column and both MA_Global and MA_Philadelphia on the row.
- I disaggregated the data by clicking Analysis and unchecking Aggregate Measure.
- I changed the chart to dual axis and then synchronized so that both lines (MA_Global and MA_Philadelphia) will have the same scale.
- I included a trend line for both lines (MA_Global and MA_Philadelphia).

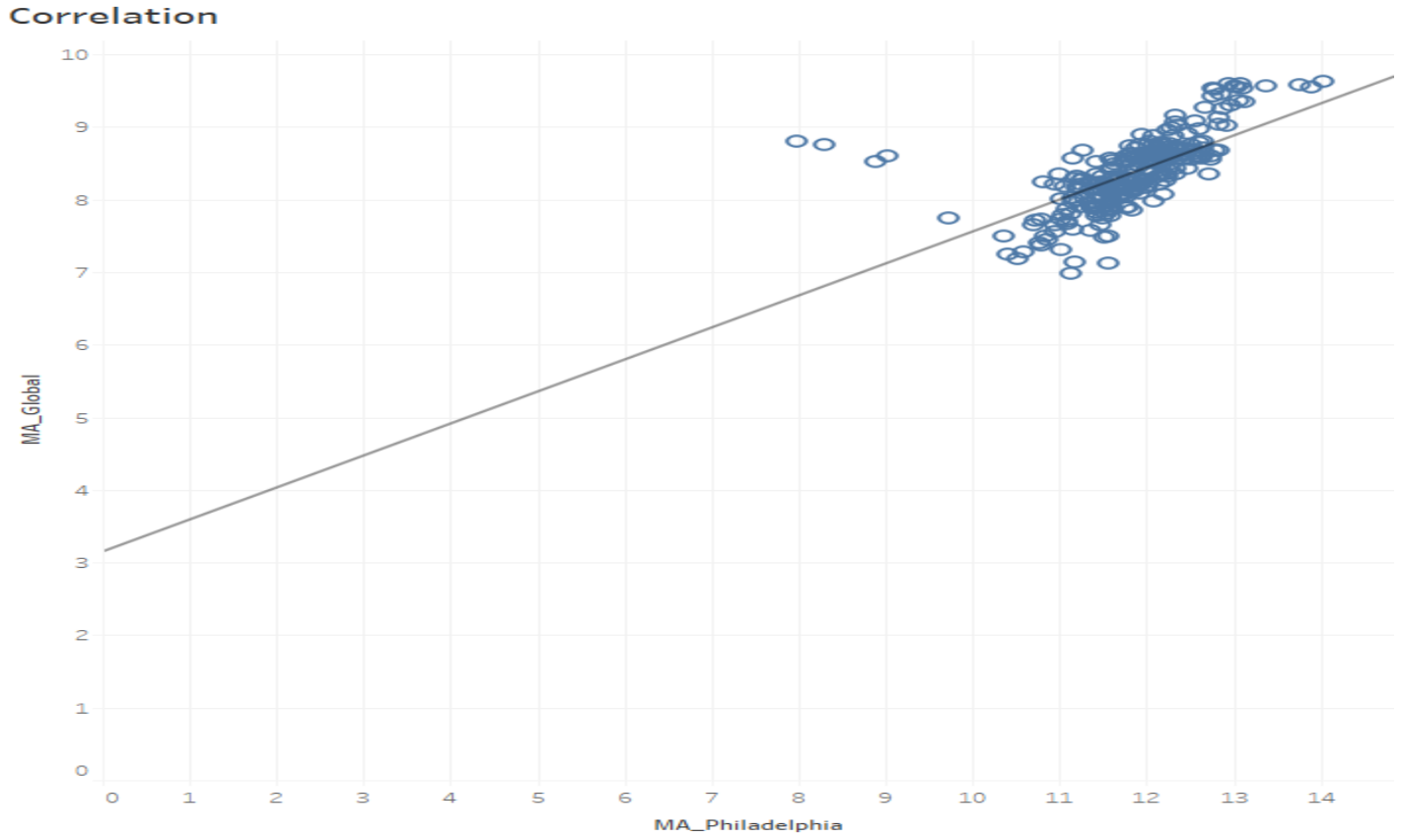
3-Year Moving Average



The trends of MA_Global and MA_Philadelphia for Year. Color shows details about MA_Global and MA_Philadelphia.

For Correlation:

- Placed MA_Global on rows and MA_Philadelphia and the columns. I added Year to the detail mark card.
- I disaggregated the data by clicking Analysis and unchecking Aggregate Measure.
- I included the trend line.
- I displayed the trend line model by right clicking on the chart, selecting trend line and selecting “Describe Trend Model.”



MA_Philadelphia vs. MA_Global.

Describe Trend Model

Trend Lines Model

A linear trend model is computed for MA_Global given MA_Philadelphia. The model may be significant at p <= 0.05.

Model formula:	(MA_Philadelphia + intercept)
Number of modeled observations:	266
Number of filtered observations:	0
Model degrees of freedom:	2
Residual degrees of freedom (DF):	264
SSE (sum squared error):	38.4583
MSE (mean squared error):	0.145675
R-Squared:	0.425521
Standard error:	0.381674
p-value (significance):	< 0.0001

Individual trend lines:

Panes		Line		Coefficients				
Row	**Column**	**p-value**	**DF**	**Term**	**Value**	**StdErr**	**t-value**	**p-value**
MA_Global	MA_Philadelphia	< 0.0001	264	MA_Philadelphia	0.440911	0.0315301	13.9838	< 0.0001
				intercept	3.15085	0.373395	8.43838	< 0.0001

Copy

Close

Insights:

1. Overall, Philadelphia has higher average temperature than the global average.
2. Starting from 1778, Philadelphia's average temperature started dropping sharply, and dropped below the global average temperature for the only time in 1780. It rose a bit in 1781 and by 1782, Philadelphia had the same average temperature as the global. Philadelphia's average temperature continued to rise steadily and never fell as low as the global temperature again.
3. Overall, both the Philadelphia and global temperatures as increasing steadily.
4. Philadelphia and global temperatures are positively correlated. R^2 of 0.42 is moderate. The model may be significant at $p<=0.05$.