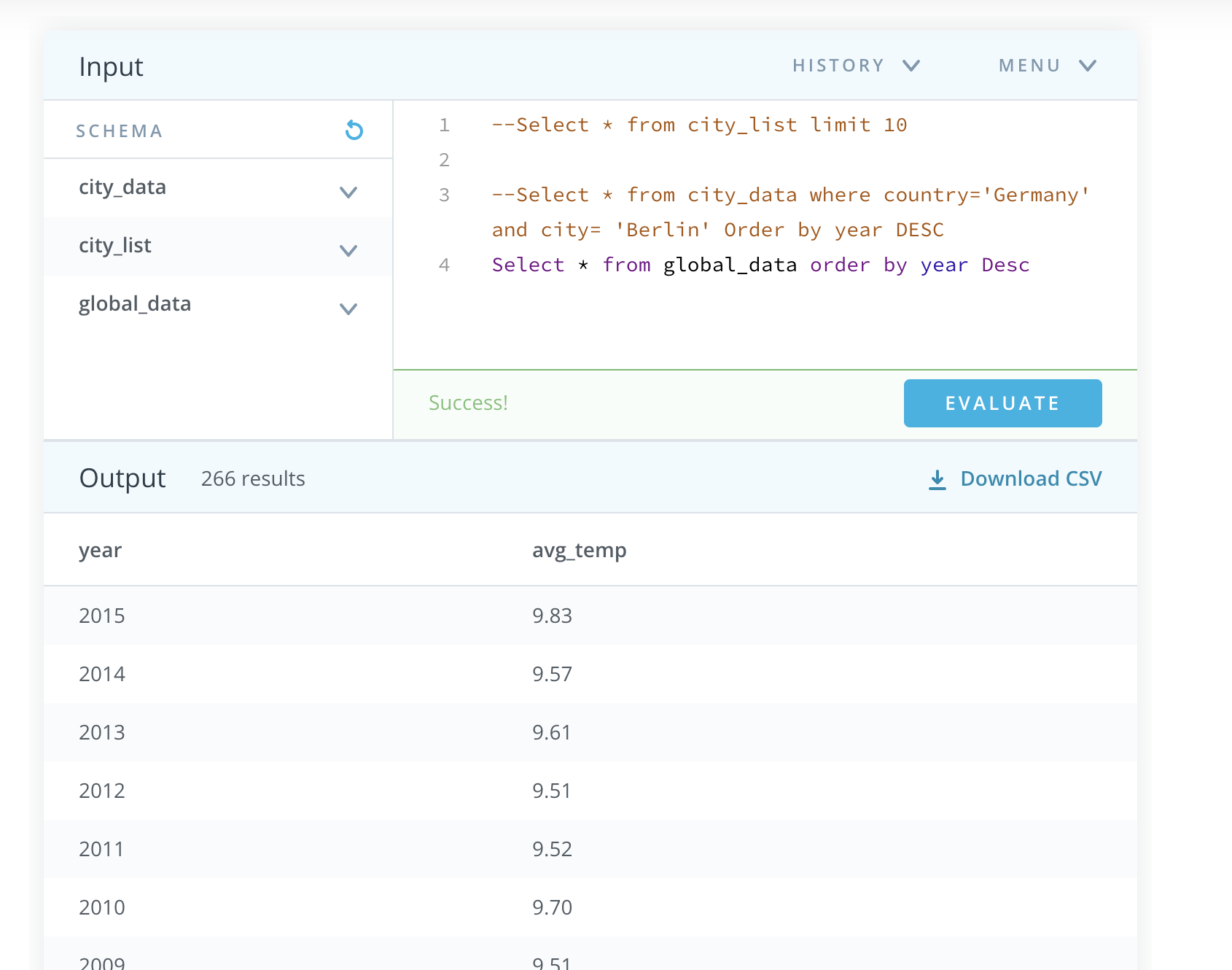
**Data Analyst Nanodegree - Project 1 - Weather trends**

**Outline**

***To start this project, I follow the instructions given by Udacity. I will enunciate this document the steps that I go through for the analysis***

***Step 1: Extract data with SQL from database in Udacity workspace***



###### **Step 2: Learn about moving averages (MA) calculation in gsheets**

[Moving average calculation in gsheet](https://docs.google.com/spreadsheets/d/1G6cKGHNO2gfi0CCs9V0w4jIvPGqY4A6Ipxqet3GXrZY/edit?usp=sharing)

***Step 3: Explore weather data in gsheets***

[Global raw data](https://docs.google.com/spreadsheets/d/1m2rn_4W-FOJPEJNvM8-oO3uI2gW_LCNiWh_0ymYOPDM/edit#gid=0)

[Local raw data](https://docs.google.com/spreadsheets/d/1m2rn_4W-FOJPEJNvM8-oO3uI2gW_LCNiWh_0ymYOPDM/edit#gid=1910417430)

***Step 4: Calculate MA in gsheets***

We look at the data since 1750 until 2013 because this is the data in common for both data sources (local and global)

Observe which moving average is the right one to look at

I chose to look at moving averages for:

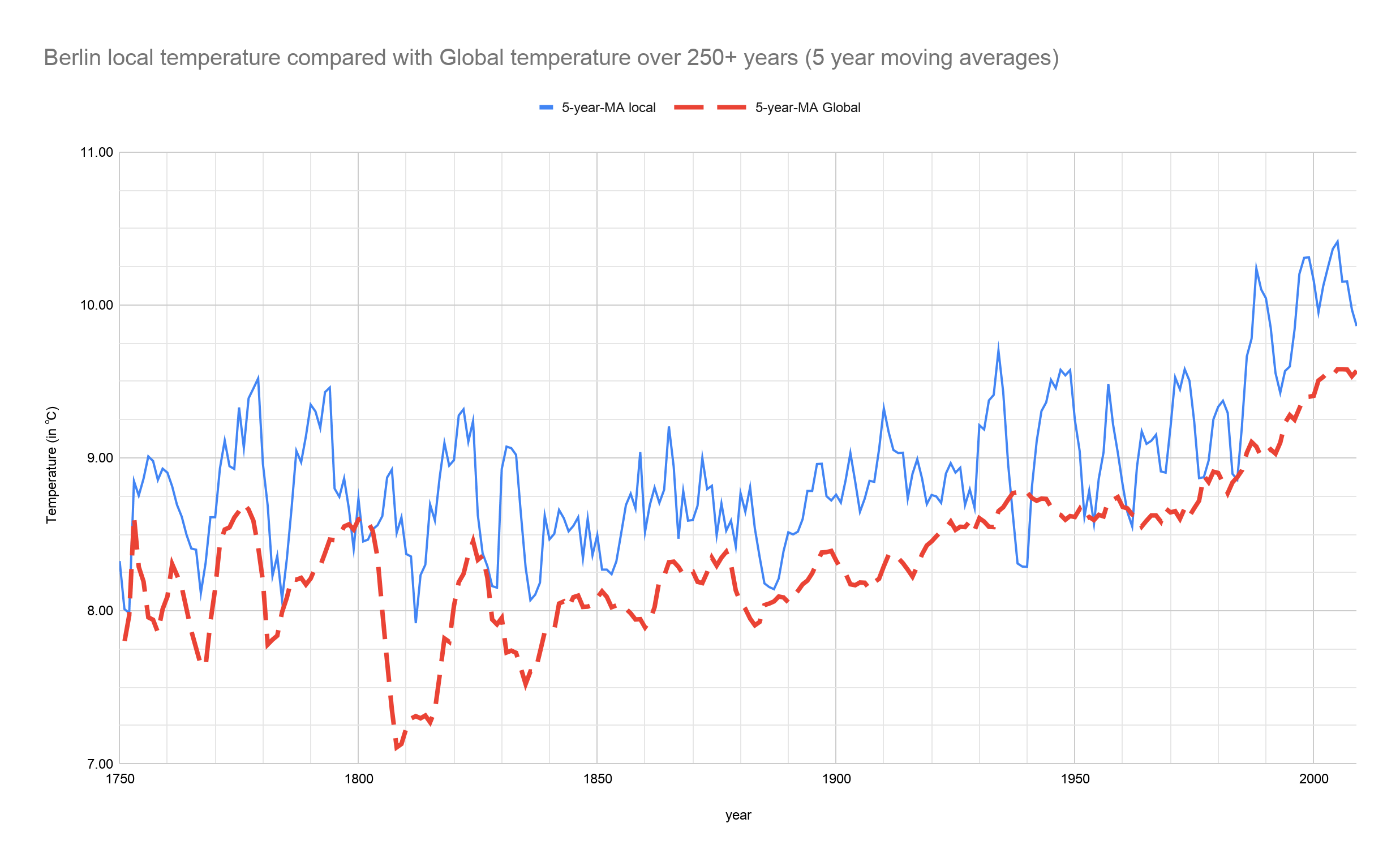
* 3 years (average of 3 years; average from 2 years before the current until current year)
* 5 years ( idem for 5 years ; average from 4 years before the current until current year)
* 7 years ( same reasoning for 7 years and so on..)
* 10 years

[Global MA](https://docs.google.com/spreadsheets/d/1m2rn_4W-FOJPEJNvM8-oO3uI2gW_LCNiWh_0ymYOPDM/edit#gid=646042096)

[Local MA](https://docs.google.com/spreadsheets/d/1m2rn_4W-FOJPEJNvM8-oO3uI2gW_LCNiWh_0ymYOPDM/edit#gid=1076677330)

After observing differences between all moving averages calculated, I opt for the 5 year MA because it is really smoothing the trend and irregular peaks that could be considered as 'exceptions' which you can still see in 3 year MA. Additionally, it doesn't seem like there are too consequential changes in a 5 year period, at the naked eye, I can see more macro temperature differences at 10 years intervals, but I would like to keep the 5 year detail in my analysis.

***Step 5: Comparison of global vs. local temperatures in gsheets***

******

[*GSheet local vs. global data comparison*](https://docs.google.com/spreadsheets/d/1m2rn_4W-FOJPEJNvM8-oO3uI2gW_LCNiWh_0ymYOPDM/edit?usp=sharing)

*Key considerations for how to visualize the trends?*

* It is a time series so it is better to visualize it as a line graph
* The vertical axis needs to be adjusted because if it starts at 0 it is difficult to observe subtle temperature variations
* Both trends are directly comparable on the same scale

***Step 6: Observations***

* + *Is your city hotter or cooler on average compared to the global average? Has the difference been consistent over time?*

Berlin is in general hotter compared with the Global temperature average and this difference has been quite consistent for the last 250 years. There were a few exceptions in history, mainly in 1798-1803 and 1936-1941 where the temperature in Berlin dropped when Global temperature was actually peaking.

* + *“How do the changes in your city’s temperatures over time compare to the changes in the global average?”*

Temperature changes in Berlin have shown high variability at period intervals of 5 years. Global temperature changes have been quite variable too but a little bit more stable. The changes in temperature in Berlin vs. Globally had a tendency to follow the same highly variable patterns before 1930 and some quite sharp drops before 1850. After that, Berlin kept this high variability over about 5 year intervals, but the global trend shows less variability. Both, though, show a clear general temperature increase from 1900's until today.

* + What does the overall trend look like? Is the world getting hotter or cooler? Has the trend been consistent over the last few hundred years?

From 1750 until 1900's, there is no clear general evolution which can be deducted. Only that, from 1850 to 1900's 10-20 years interval average temperatures tend to be higher than before. However, when we arrive from 1900 to 1950's, global trend is clearly showing an increase. It is more difficult to observe in Berlin because the variability as remained high but still on average compared with 100 years before that you can see it is on average higher temperatures. From 1950 on tot today, there is no doubt for both local and global temperatures that there is a general increase.