Data Analysis Nanodegree project 1

In this project I use the data provided from Udacity which was whether data. First, I chose a city of my own and its average temperatures and then compare it with the global temperatures. The period I took was from 1847 to 2013. Secondly, I opened the csv files and then join the city and the global data together. Third, I calculated the moving average using excel with a 5-year period. Lastly, I did a line chart visualization.

Tools and Programming used in this project:

1. SQL – Extract data
2. Excel – clean, calculate moving average, and visualization.

Now calculating the moving average was very simple using Excel. I used periods separated by 5 years, I used the data analysis tool in Excel which really did most of the work. Here is an example of the csv file.

	A	B	C	D	E	F
1	year	city	Riyadh_temp	Global_temp	Riyadh_MA	Global_MA
2	1847	Riyadh	25.3	8.09	25.3	8.09
3	1848	Riyadh	24.56	7.98	24.93	8.035
4	1849	Riyadh	24.8	7.98	24.89	8.017
5	1850	Riyadh	24.34	7.9	24.75	7.9875
6	1851	Riyadh	25.03	8.18	24.806	8.026
7	1852	Riyadh	24.85	8.1	24.716	8.028
8	1853	Riyadh	24.93	8.04	24.79	8.04
9	1854	Riyadh	24.72	8.21	24.774	8.086
10	1855	Riyadh	24.92	8.11	24.89	8.128
11	1856	Riyadh	24.57	8	24.798	8.092
12	1857	Riyadh	24.26	7.76	24.68	8.024
13	1858	Riyadh	25.01	8.1	24.696	8.036
14	1859	Riyadh	24.95	8.25	24.742	8.044
15	1860	Riyadh	24.94	7.96	24.746	8.014
16	1861	Riyadh	24.13	7.85	24.658	7.984
17	1862	Riyadh	23.77	7.56	24.56	7.944
18	1863	Riyadh	24.28	8.11	24.414	7.946
19	1864	Riyadh	25.03	7.98	24.43	7.892
20	1865	Riyadh	25.23	8.18	24.488	7.936
21	1866	Riyadh	24.92	8.29	24.646	8.024
22	1867	Riyadh	25.22	8.44	24.936	8.2
23	1868	Riyadh	25	8.25	25.08	8.228
24	1869	Riyadh	25.3	8.43	25.134	8.318
25	1870	Riyadh	25.02	8.2	25.092	8.322
26	1871	Riyadh	24.73	8.12	25.054	8.288
27	1872	Riyadh	24.87	8.19	24.984	8.238
28	1873	Riyadh	25.24	8.35	25.032	8.258
29	1874	Riyadh	24.98	8.43	24.968	8.258
30	1875	Riyadh	24.43	7.86	24.85	8.19
31	1876	Riyadh	24.89	8.08	24.882	8.182

I used a line chart because that is the appropriate visual for visualizing years and time.

2 SQL Queries, one for my city and one for the global data.

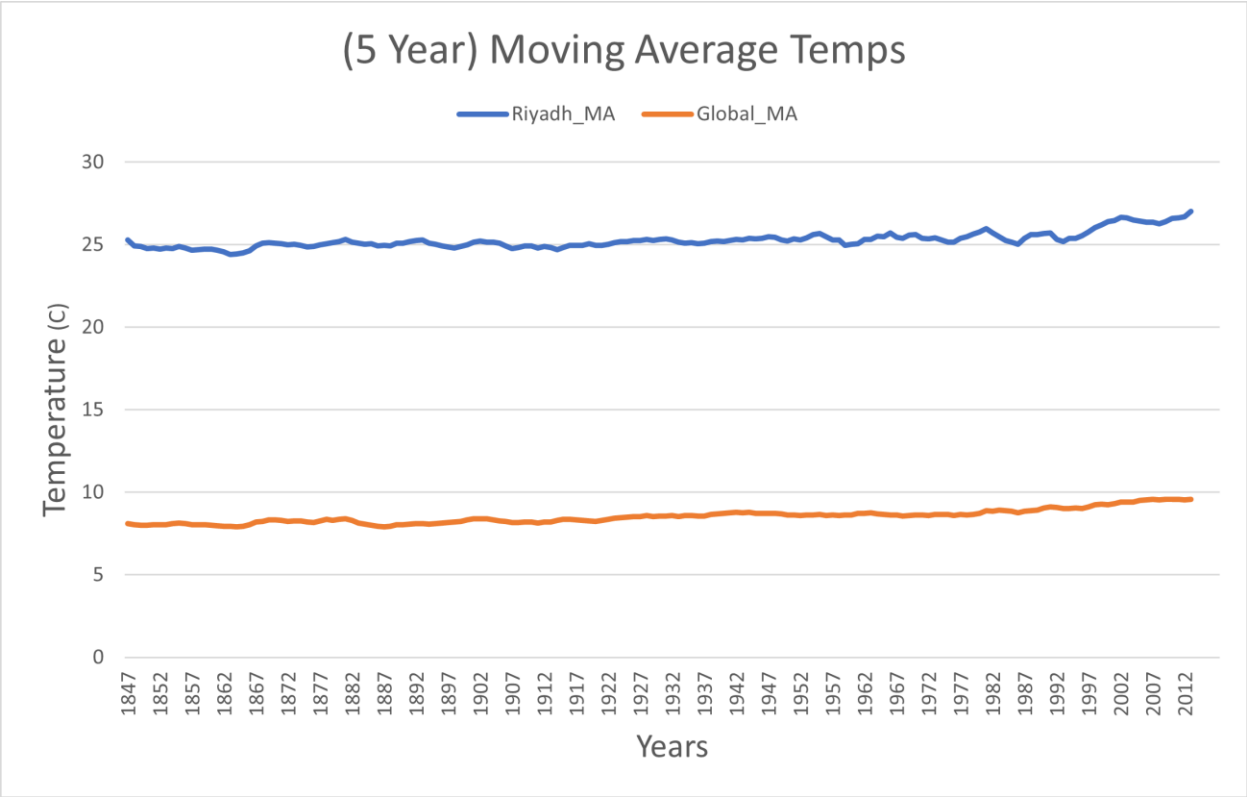
The global data query:

- ```
  SELECT *
  FROM global_data
  ```

My city's query:

- ```
  SELECT city, year, avg_temp
  FROM city_data
  WHERE city='Riyadh';
  ```

This is the Line chart:



Observations:

1. It is obvious that Riyadh's Average temperatures is hotter than the global average temperatures
2. In 1992 and after we can see that Riyadh's temperature is increasing.
3. The global average temperatures are colder than Riyadh.
4. We can say that both Riyadh and the global temperatures are all increasing with time.