Project 1 for Udacity program
Data analysis Nanodegree
Explore weather trends

By:Fahad AlQahtani

4/25/2020

Index

Introduction:

This is the first project for my data analysis Nanodegree, In this project, I will analyze local and global temperature data and compare the temperature trends where I live to overall global temperature trends, In this project I will get the dataset form the database by using SQL QUERY, Then I will do some analysis to this data for example get the average for temperature, MAX, etc.. , Then visualize it into charts by using Microsoft excel.

Extract the data:

In this section I will explain how I get the dataset from the database.

1.In core curriculum section there is “Welcome to the Nanodegree” when I click on it and finish all the lessons I go through project, In this project there is concept called “Accessing data with SQL”.

2.When I click it and they ask me to write a query to get the data from the database and download it as “.CSV” file .

3.I wrote the query and get the data from database and download it as “.CSV” file

3.This the query I wrote to get the data:

(SELECT gd.year, cd.country, cd.city,

cd.avg_temp Riyadh, gd.avg_temp World

FROM city_data cd JOIN global_data gd ON gd.year = cd.year

WHERE cd.city = 'Riyadh'

ORDER BY gd.year)

- (SELECT gd.year, cd.country, cd.city, cd.avg_temp Riyadh, gd.avg_temp World) :

    - ❖ This mean select gd.year and gd.avg_temp(and name the column “World”) from GD Table.
    - ❖ This also measn select cd.country and cd.city , cd._avg_temp(and name the column “Riyadh”) from CD table.

- (FROM city_data cd JOIN global_data gd ON gd.year = cd.year) :

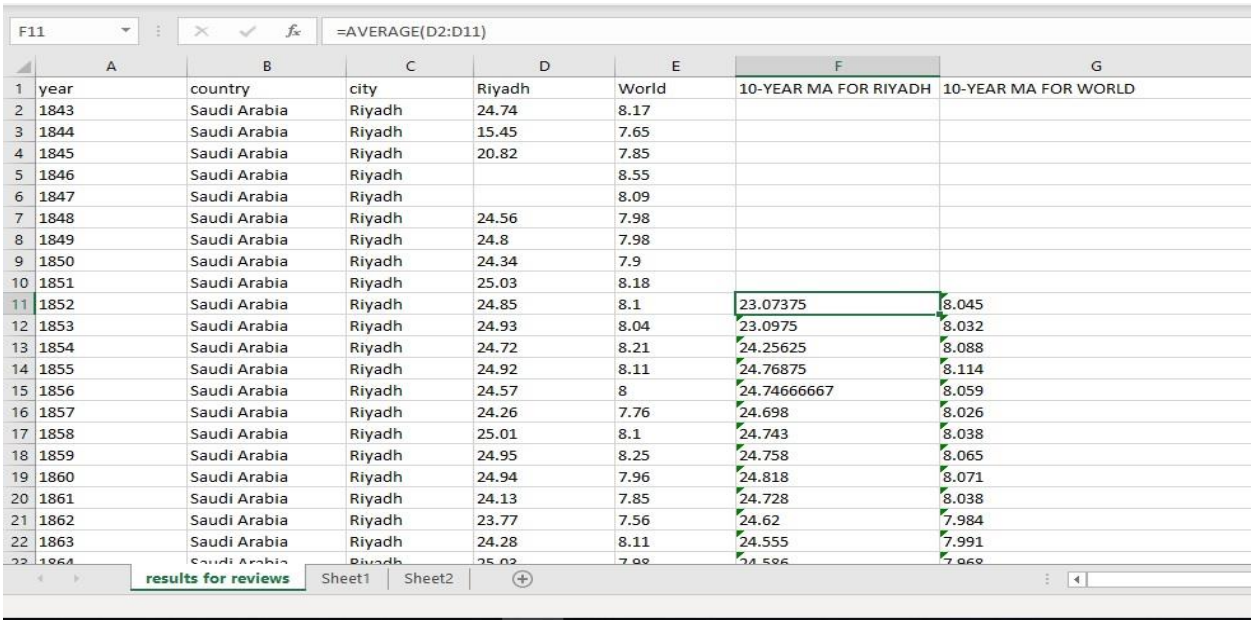
    - ❖ This mean what we select is from these both tables city_data and global_data and make a abbreviation for both of it (CD,GD) to call it when you need it easier and faster.
    - ❖ JOIN mean make relationship between these 2 tables on column named (YEAR) on both tables.

- (WHERE cd.city = 'Riyadh') :
    - ❖ This mean filter the cites and bring Riyadh only form city_data .
- (ORDER BY gd.year):
    - ❖ This mean sort the data by year.

Open up the CSV:

In this section I will explain what I did with this CSV files by using excel.

1.After I extract this dataset, Then I create a “Moving average” for every 10 years for both (World, Riyadh) as shown in this figure for Riyadh.

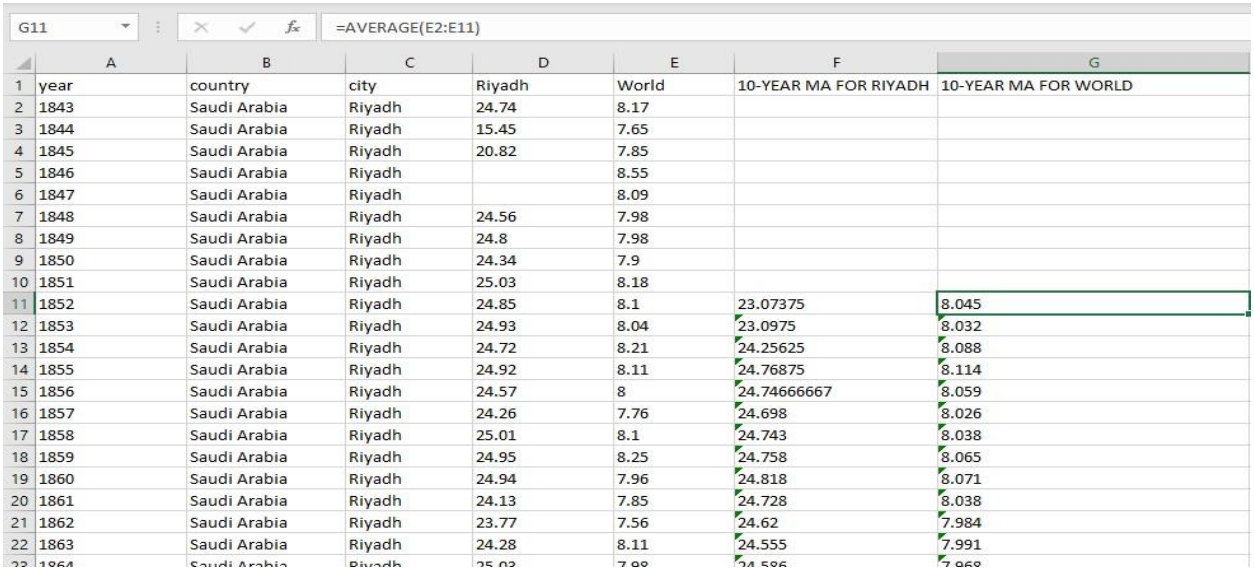


F11 =AVERAGE(D2:D11)

	A	B	C	D	E	F	G
1	year	country	city	Riyadh	World	10-YEAR MA FOR RIYADH	10-YEAR MA FOR WORLD
2	1843	Saudi Arabia	Riyadh	24.74	8.17		
3	1844	Saudi Arabia	Riyadh	15.45	7.65		
4	1845	Saudi Arabia	Riyadh	20.82	7.85		
5	1846	Saudi Arabia	Riyadh		8.55		
6	1847	Saudi Arabia	Riyadh		8.09		
7	1848	Saudi Arabia	Riyadh	24.56	7.98		
8	1849	Saudi Arabia	Riyadh	24.8	7.98		
9	1850	Saudi Arabia	Riyadh	24.34	7.9		
10	1851	Saudi Arabia	Riyadh	25.03	8.18		
11	1852	Saudi Arabia	Riyadh	24.85	8.1	23.07375	8.045
12	1853	Saudi Arabia	Riyadh	24.93	8.04	23.0975	8.032
13	1854	Saudi Arabia	Riyadh	24.72	8.21	24.25625	8.088
14	1855	Saudi Arabia	Riyadh	24.92	8.11	24.76875	8.114
15	1856	Saudi Arabia	Riyadh	24.57	8	24.74666667	8.059
16	1857	Saudi Arabia	Riyadh	24.26	7.76	24.698	8.026
17	1858	Saudi Arabia	Riyadh	25.01	8.1	24.743	8.038
18	1859	Saudi Arabia	Riyadh	24.95	8.25	24.758	8.065
19	1860	Saudi Arabia	Riyadh	24.94	7.96	24.818	8.071
20	1861	Saudi Arabia	Riyadh	24.13	7.85	24.728	8.038
21	1862	Saudi Arabia	Riyadh	23.77	7.56	24.62	7.984
22	1863	Saudi Arabia	Riyadh	24.28	8.11	24.555	7.991
23	1864	Saudi Arabia	Riyadh	25.03	7.88	24.586	7.968

results for reviews | Sheet1 | Sheet2

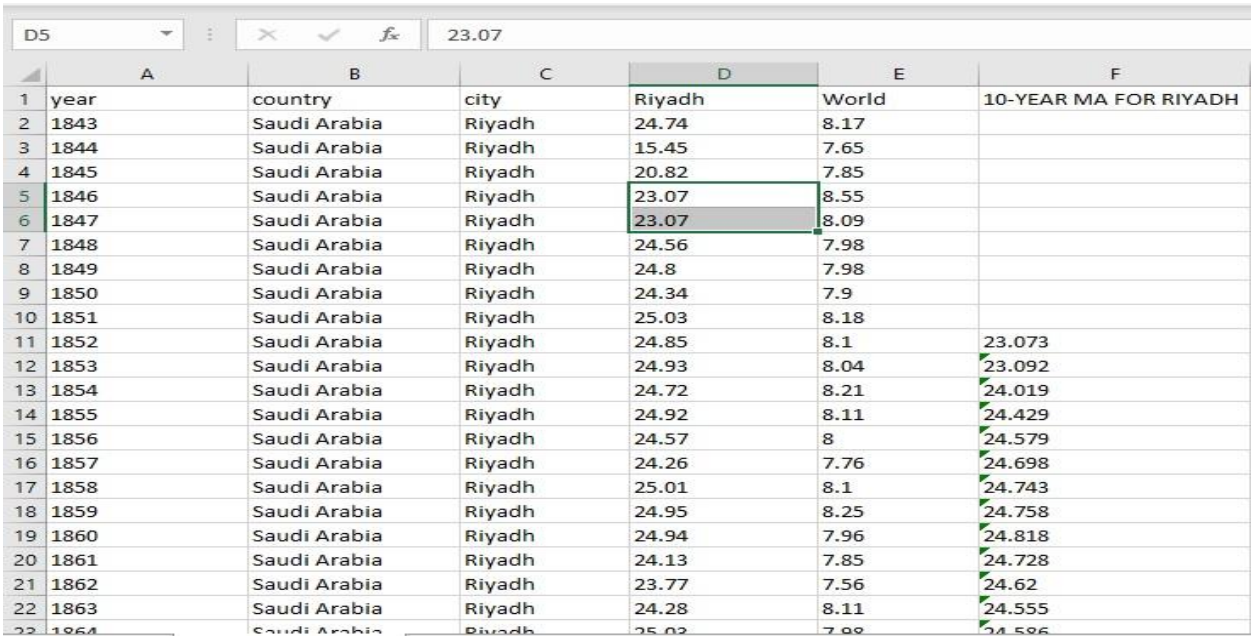
And for the world



G11 =AVERAGE(E2:E11)

	A	B	C	D	E	F	G
1	year	country	city	Riyadh	World	10-YEAR MA FOR RIYADH	10-YEAR MA FOR WORLD
2	1843	Saudi Arabia	Riyadh	24.74	8.17		
3	1844	Saudi Arabia	Riyadh	15.45	7.65		
4	1845	Saudi Arabia	Riyadh	20.82	7.85		
5	1846	Saudi Arabia	Riyadh		8.55		
6	1847	Saudi Arabia	Riyadh		8.09		
7	1848	Saudi Arabia	Riyadh	24.56	7.98		
8	1849	Saudi Arabia	Riyadh	24.8	7.98		
9	1850	Saudi Arabia	Riyadh	24.34	7.9		
10	1851	Saudi Arabia	Riyadh	25.03	8.18		
11	1852	Saudi Arabia	Riyadh	24.85	8.1	23.07375	8.045
12	1853	Saudi Arabia	Riyadh	24.93	8.04	23.0975	8.032
13	1854	Saudi Arabia	Riyadh	24.72	8.21	24.25625	8.088
14	1855	Saudi Arabia	Riyadh	24.92	8.11	24.76875	8.114
15	1856	Saudi Arabia	Riyadh	24.57	8	24.74666667	8.059
16	1857	Saudi Arabia	Riyadh	24.26	7.76	24.698	8.026
17	1858	Saudi Arabia	Riyadh	25.01	8.1	24.743	8.038
18	1859	Saudi Arabia	Riyadh	24.95	8.25	24.758	8.065
19	1860	Saudi Arabia	Riyadh	24.94	7.96	24.818	8.071
20	1861	Saudi Arabia	Riyadh	24.13	7.85	24.728	8.038
21	1862	Saudi Arabia	Riyadh	23.77	7.56	24.62	7.984
22	1863	Saudi Arabia	Riyadh	24.28	8.11	24.555	7.991
23	1864	Saudi Arabia	Riyadh	25.03	7.88	24.586	7.968

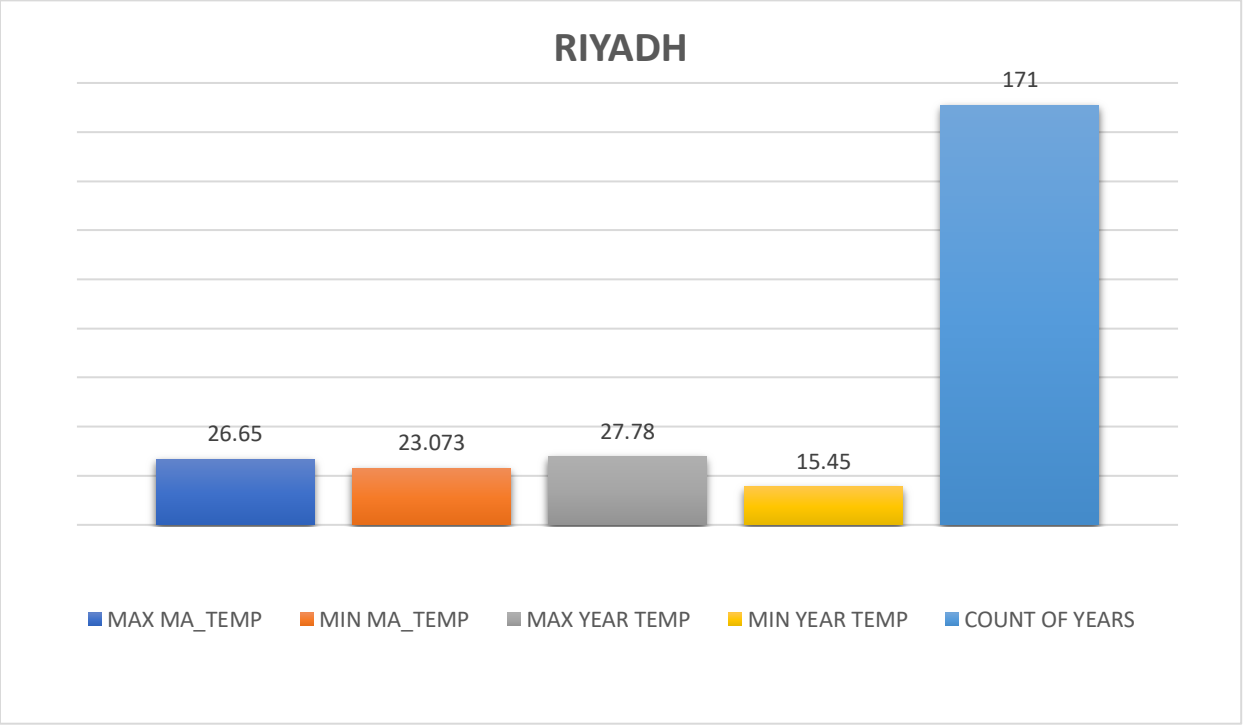
2.I saw there is a missing value for avg_temp in 2 years (1846, 1847) in Riyadh, and I replace it by there own 10-year MA as shown in the next figure.

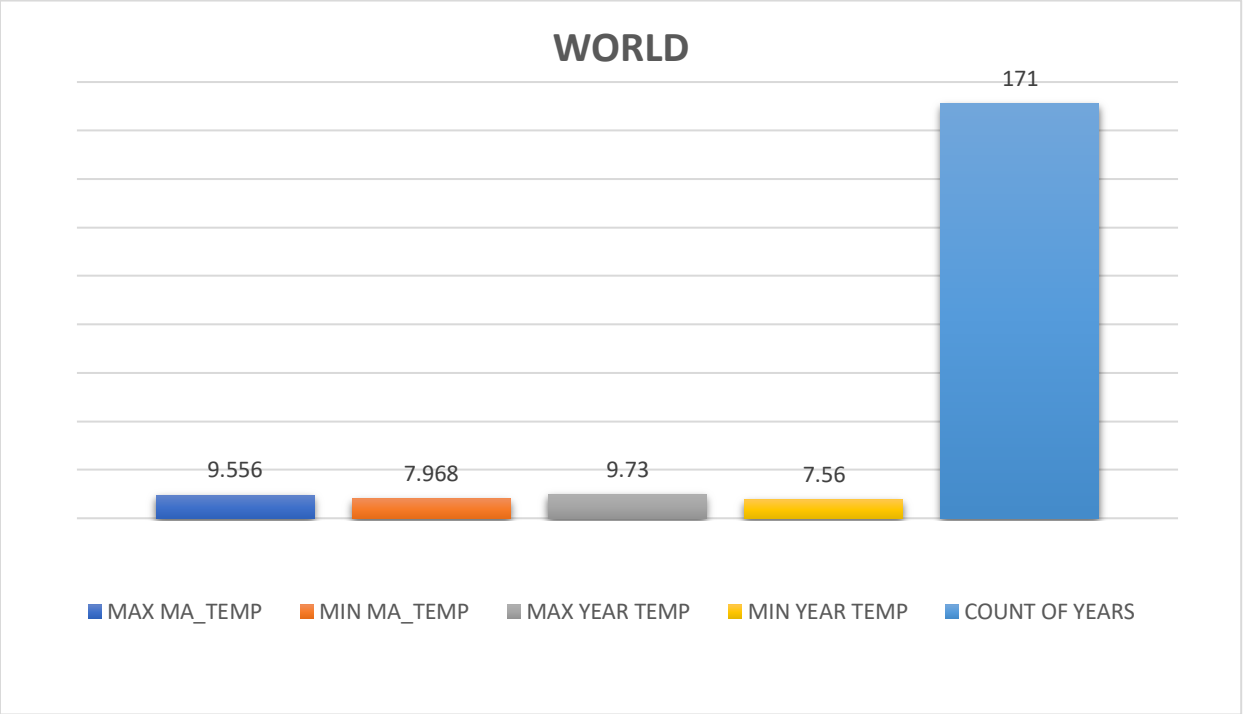


D5 | 23.07

	A	B	C	D	E	F
1	year	country	city	Riyadh	World	10-YEAR MA FOR RIYADH
2	1843	Saudi Arabia	Riyadh	24.74	8.17	
3	1844	Saudi Arabia	Riyadh	15.45	7.65	
4	1845	Saudi Arabia	Riyadh	20.82	7.85	
5	1846	Saudi Arabia	Riyadh	23.07	8.55	
6	1847	Saudi Arabia	Riyadh	23.07	8.09	
7	1848	Saudi Arabia	Riyadh	24.56	7.98	
8	1849	Saudi Arabia	Riyadh	24.8	7.98	
9	1850	Saudi Arabia	Riyadh	24.34	7.9	
10	1851	Saudi Arabia	Riyadh	25.03	8.18	
11	1852	Saudi Arabia	Riyadh	24.85	8.1	23.073
12	1853	Saudi Arabia	Riyadh	24.93	8.04	23.092
13	1854	Saudi Arabia	Riyadh	24.72	8.21	24.019
14	1855	Saudi Arabia	Riyadh	24.92	8.11	24.429
15	1856	Saudi Arabia	Riyadh	24.57	8	24.579
16	1857	Saudi Arabia	Riyadh	24.26	7.76	24.698
17	1858	Saudi Arabia	Riyadh	25.01	8.1	24.743
18	1859	Saudi Arabia	Riyadh	24.95	8.25	24.758
19	1860	Saudi Arabia	Riyadh	24.94	7.96	24.818
20	1861	Saudi Arabia	Riyadh	24.13	7.85	24.728
21	1862	Saudi Arabia	Riyadh	23.77	7.56	24.62
22	1863	Saudi Arabia	Riyadh	24.28	8.11	24.555
23	1864	Saudi Arabia	Riyadh	25.03	7.98	24.586

3.Then I take the average and MAX, MIN for all the avg_temp, and the same for 10-Year MA for both and visualize it.





Create a line chart:

In this section I will explain how I create line chart to compare between Riyadh and World by using excel.

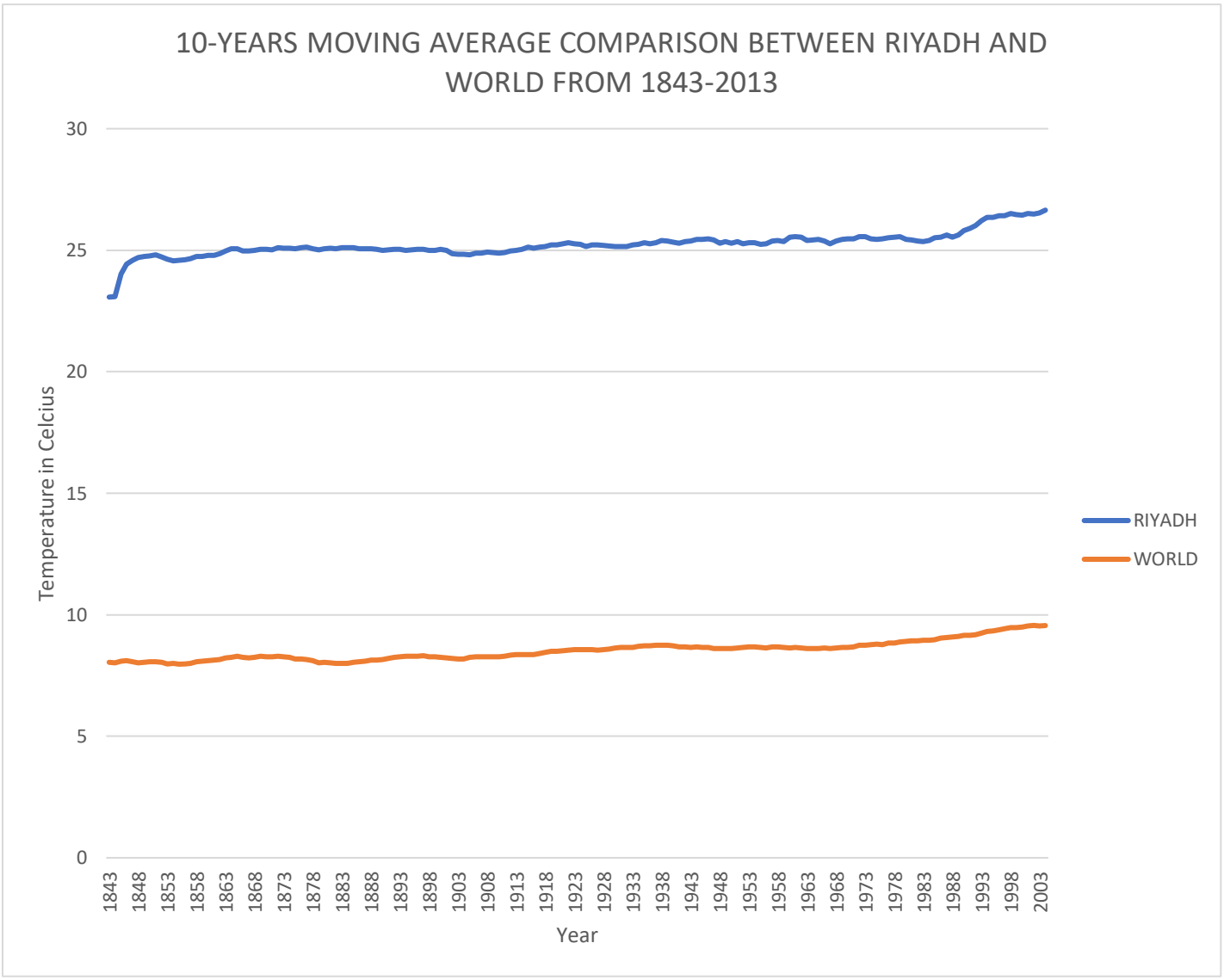
1.I selected the data from Riyadh and create the line chart after that I add the data from the World by click on select data.

2.I create series 1 for Riyadh and name it Riyadh and the second for World .

3.After that I edit series and select 10-year MA for series value for both.

4.In x-axis label I use years to display the 10-year MA through it.

5.This is the line chart:



Observations:

In this section I will write my observations of Riyadh and the World temperature.

- As shown in the line chart Riyadh is hotter than the World.
- As shown in the line chart over the years the temperature in all the world is increasing and this mean the world well be hotter in the next years.
- As shown in the 2 columns charts:
  - ❖ Riyadh MAX average over all years is 27.78 and its in 2013, And the MIN is 15.45 and its in 1844.
  - ❖ Riyadh MAX 10-Year MA is 26.65 and its from 2004 to 2013, And the MIN is 23.073 and its from 1843 to 1852.

(These 2 points explain is the temperature increasing over the years , and also you can see the max and min years is included to the 10-year MA).

- ❖ World MAX average over all years is 9.73 and its in 2007, And the MIN is 7.56 and its in 1862.
- ❖ Global MAX 10-Year MA is 9.556 and its from 2004 to 2013, And the MIN is 7.968 and its from 1855 to 1864.

(These 2 points explain is the temperature increasing over the years , and also you can see the max and min years is included to the 10-year MA but Riyadh First 10 years is the coldest and World between The second and third 10 years is the coldest,

Also both Riyadh and The World has same hottest 10 years and it's the last 10 years.).

- Riyadh temperature is increased more than the global.
- Riyadh temperature is more than global at least 17C degree.
- Over The 10-year MA from (1843-1852) to (2004-2013) Riyadh increased by 3.577 C degree, and the global increased by 1.511 C degree.