Assignment 1: Exploring Weather Trends - Project

1. Extract data

   I pulled out all of the dataset by using SQL:

   a. City_data:

   SELECT * FROM city_data

   b. City_list

   SELECT * FROM city_list

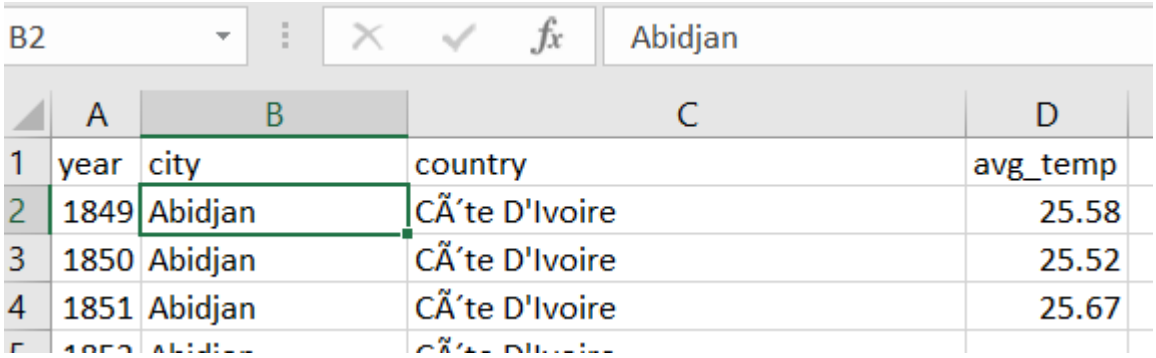
   c. Global_data

   SELECT * FROM global_data

2. Open up the CSV

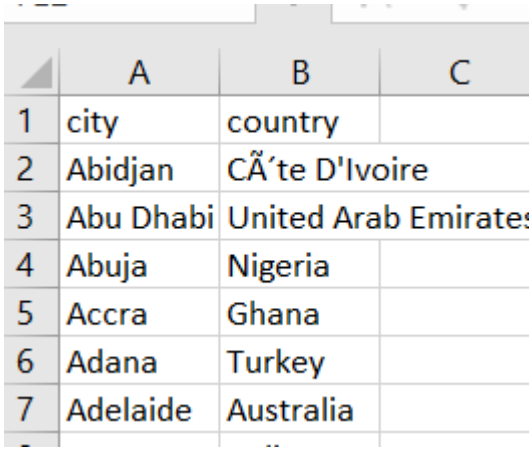
   I download the result from previous task to achieve 3 .csv files. I manage to open the up them in excel files:

   a. The result of city_data: an excel file with 4 columns – year, city, country, and avg_temp for every year in each city which the list is in city_list file



	A	B	C	D
1	year	city	country	avg_temp
2	1849	Abidjan	CÃ´te D'Ivoire	25.58
3	1850	Abidjan	CÃ´te D'Ivoire	25.52
4	1851	Abidjan	CÃ´te D'Ivoire	25.67

   b. City_list file: An excel file with 2 columns – city and country correspond with this city



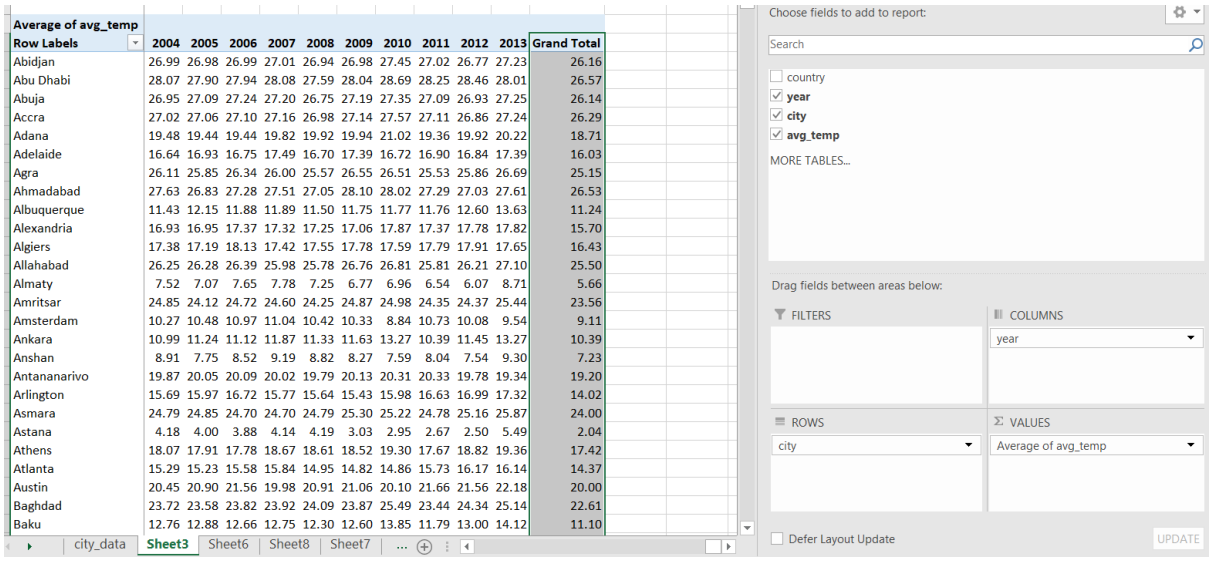
	A	B	C
1	city	country	
2	Abidjan	CÃ´te D'Ivoire	
3	Abu Dhabi	United Arab Emirates	
4	Abuja	Nigeria	
5	Accra	Ghana	
6	Adana	Turkey	
7	Adelaide	Australia	

   c. Global_data: an excel file shows the global temperature's observation each year, from 1750 o 2015

	A	B	C
1	year	avg_temp	
2	1750	8.72	
3	1751	7.98	
4	1752	5.78	
5	1753	8.39	
6	1754	8.47	
7	1755	8.36	
8	1756	8.85	
9	1757	9.02	
10	1758	6.74	

d. Select the cities to compare the temperature: I am living in Perth, Australia. So, it should be my fist selection. I was born in Hochiminh city. So, it pick it up as the second entities. I will choose 2 others by querying the city which the overall evg_temp are highest and lowest.

From the city_data sheet, I generate a pivot table with the row is city name, columns are each observation year and the grand total is average of avg_temp in the city per number of observation's year.



Average of avg_temp											
Row Labels	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Grand Total
Abidjan	26.99	26.98	26.99	27.01	26.94	26.98	27.45	27.02	26.77	27.23	26.16
Abu Dhabi	28.07	27.90	27.94	28.08	27.59	28.04	28.69	28.25	28.46	28.01	26.57
Abuja	26.95	27.09	27.24	27.20	26.75	27.19	27.35	27.09	26.93	27.25	26.14
Accra	27.02	27.06	27.10	27.16	26.98	27.14	27.57	27.11	26.86	27.24	26.29
Adana	19.48	19.44	19.44	19.82	19.92	19.94	21.02	19.36	19.92	20.22	18.71
Adelaide	16.64	16.93	16.75	17.49	16.70	17.39	16.72	16.90	16.84	17.39	16.03
Agra	26.11	25.85	26.34	26.00	25.57	26.55	26.51	25.53	25.86	26.69	25.15
Ahmadabad	27.63	26.83	27.28	27.51	27.05	28.10	28.02	27.29	27.03	27.61	26.53
Albuquerque	11.43	12.15	11.88	11.89	11.50	11.75	11.77	11.76	12.60	13.63	11.24
Alexandria	16.93	16.95	17.37	17.32	17.25	17.06	17.87	17.37	17.78	17.82	15.70
Algiers	17.38	17.19	18.13	17.42	17.55	17.78	17.59	17.79	17.91	17.65	16.43
Allahabad	26.25	26.28	26.39	25.98	25.78	26.76	26.81	25.81	26.21	27.10	25.50
Almaty	7.52	7.07	7.65	7.78	7.25	6.77	6.96	6.54	6.07	8.71	5.66
Amritsar	24.85	24.12	24.72	24.60	24.25	24.87	24.98	24.35	24.37	25.44	23.56
Amsterdam	10.27	10.48	10.97	11.04	10.42	10.33	8.84	10.73	10.08	9.54	9.11
Ankara	10.99	11.24	11.12	11.87	11.33	11.63	13.27	10.39	11.45	13.27	10.39
Anshan	8.91	7.75	8.52	9.19	8.82	8.27	7.59	8.04	7.54	9.30	7.23
Antananarivo	19.87	20.05	20.09	20.02	19.79	20.13	20.31	20.33	19.78	19.34	19.20
Arlington	15.69	15.97	16.72	15.77	15.64	15.43	15.98	16.63	16.99	17.32	14.02
Asmara	24.79	24.85	24.70	24.70	24.79	25.30	25.22	24.78	25.16	25.87	24.00
Astana	4.18	4.00	3.88	4.14	4.19	3.03	2.95	2.67	2.50	5.49	2.04
Athens	18.07	17.91	17.78	18.67	18.61	18.52	19.30	17.67	18.82	19.36	17.42
Atlanta	15.29	15.23	15.58	15.84	14.95	14.82	14.86	15.73	16.17	16.14	14.37
Austin	20.45	20.90	21.56	19.98	20.91	21.06	20.10	21.66	21.56	22.18	20.00
Baghdad	23.72	23.58	23.82	23.92	24.09	23.87	25.49	23.44	24.34	25.14	22.61
Baku	12.76	12.88	12.66	12.75	12.30	12.60	13.85	11.79	13.00	14.12	11.10

From the Grand Total column, which indicate the average of agv_temp in each city, I extract the highest and lowest values and city corresponded with those values. So, the final selection to be appeared as list:

Selected cities compared with global data		
City	**Total Count of avg_temp**	**Total Average of avg_temp**
Ulaanbaatar	72.93%	-3.37
Khartoum	56.02%	29.06

Ho Chi Minh City	65.41%	27.18
Perth	53.38%	18.26
Global	100.00%	

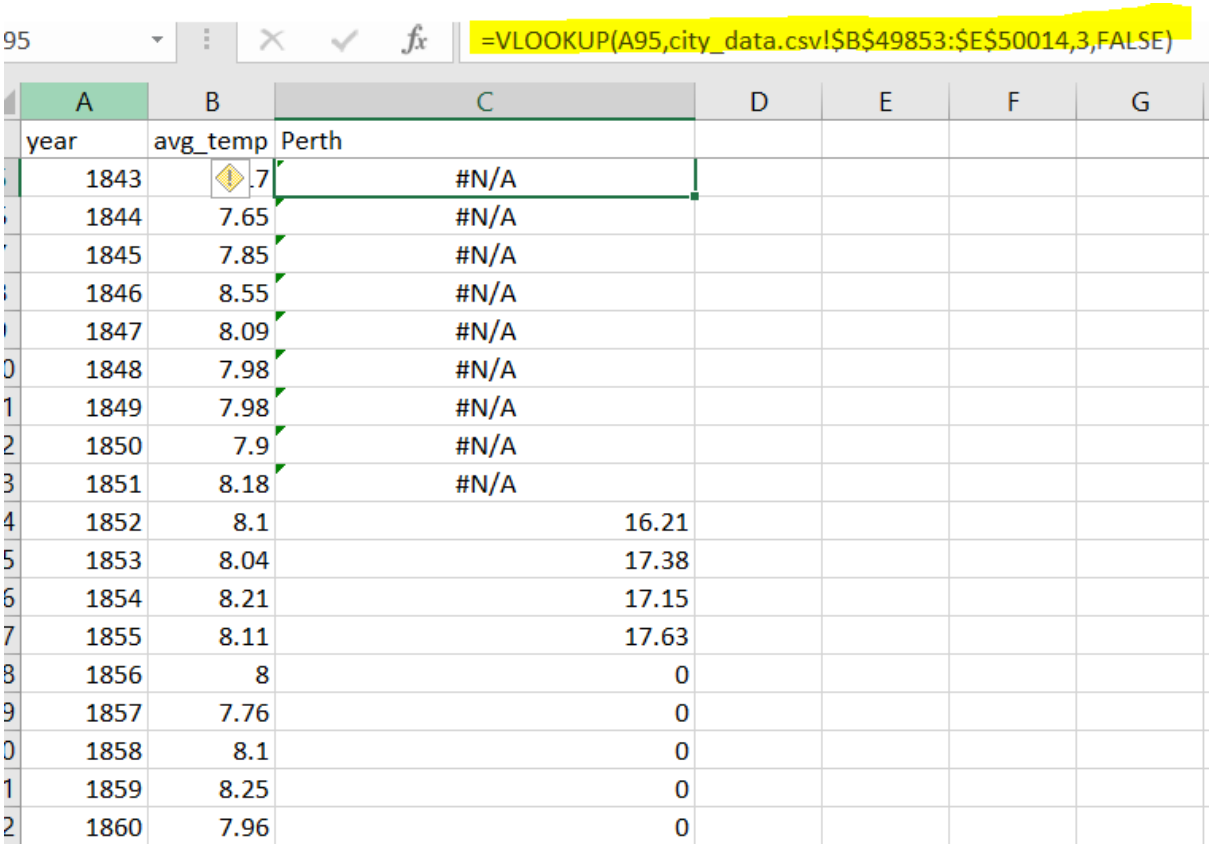
The table shows the name of 4 selected countries, indicate by total of non-missing values (the percentage of years the value is not missing compared with total observation's years which taken in global data), total Counts of avg_temp, and overral average of avg_temp

3. Create the line chart
    a. Merge city_data table with global_data

    From city_data excel -> Data -> Filter ->Select Perth from list of city name, we got the filter table of Perth's temperature data over the observation years

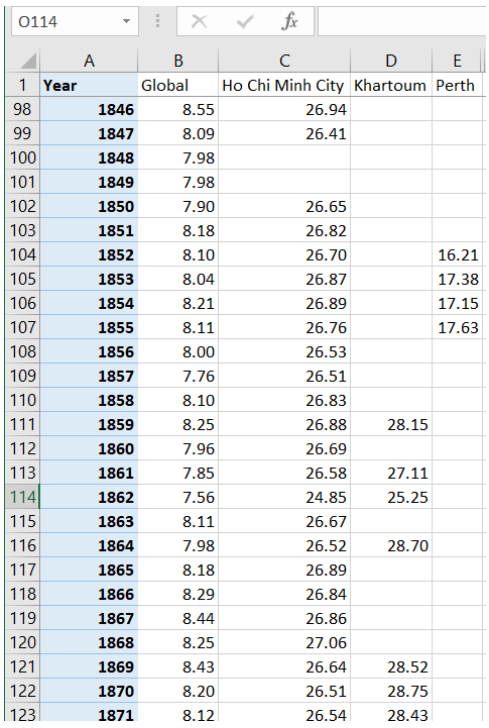
A	B	C	D
country	year	city	avg_ter
Australia	1852	Perth	16.21
Australia	1853	Perth	17.38
Australia	1854	Perth	17.15
Australia	1855	Perth	17.63
Australia	1856	Perth	
Australia	1857	Perth	
Australia	1858	Perth	
Australia	1859	Perth	
Australia	1860	Perth	
Australia	1861	Perth	
Australia	1862	Perth	
Australia	1863	Perth	
Australia	1864	Perth	
Australia	1865	Perth	
Australia	1866	Perth	
Australia	1867	Perth	
Australia	1868	Perth	
Australia	1869	Perth	
Australia	1870	Perth	
Australia	1871	Perth	
Australia	1872	Perth	
Australia	1873	Perth	
Australia	1874	Perth	
Australia	1875	Perth	
Australia	1876	Perth	17.58
Australia	1877	Perth	18.16

From global_data, using VLOOKUP function to query temperature values of Perth based on observation years. The missing value would be returned by #N/A value. This values need to be eliminated by Find #N/A and replace by blank



	A	B	C	D	E	F	G
	year	avg_temp	Perth				
	1843	.7	#N/A				
	1844	7.65	#N/A				
	1845	7.85	#N/A				
	1846	8.55	#N/A				
	1847	8.09	#N/A				
0	1848	7.98	#N/A				
1	1849	7.98	#N/A				
2	1850	7.9	#N/A				
3	1851	8.18	#N/A				
4	1852	8.1	16.21				
5	1853	8.04	17.38				
6	1854	8.21	17.15				
7	1855	8.11	17.63				
8	1856	8	0				
9	1857	7.76	0				
0	1858	8.1	0				
1	1859	8.25	0				
2	1860	7.96	0				

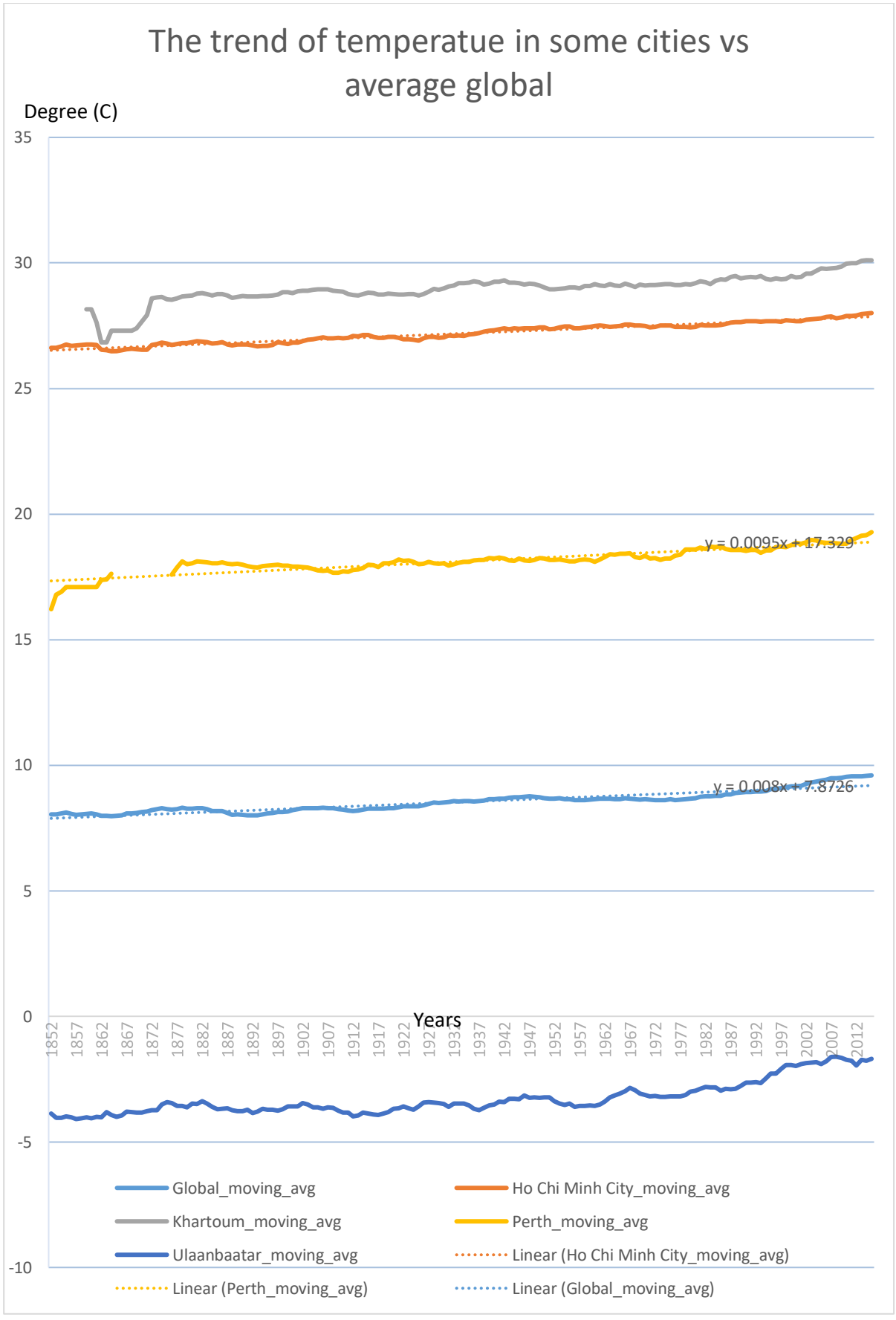
Do the same for others cities, the results we got is the temperature table of selected city compared with global_data:



	A	B	C	D	E
1	**Year**	Global	Ho Chi Minh City	Khartoum	Perth
98	**1846**	8.55	26.94		
99	**1847**	8.09	26.41		
100	**1848**	7.98			
101	**1849**	7.98			
102	**1850**	7.90	26.65		
103	**1851**	8.18	26.82		
104	**1852**	8.10	26.70		16.21
105	**1853**	8.04	26.87		17.38
106	**1854**	8.21	26.89		17.15
107	**1855**	8.11	26.76		17.63
108	**1856**	8.00	26.53		
109	**1857**	7.76	26.51		
110	**1858**	8.10	26.83		
111	**1859**	8.25	26.88	28.15	
112	**1860**	7.96	26.69		
113	**1861**	7.85	26.58	27.11	
114	**1862**	7.56	24.85	25.25	
115	**1863**	8.11	26.67		
116	**1864**	7.98	26.52	28.70	
117	**1865**	8.18	26.89		
118	**1866**	8.29	26.84		
119	**1867**	8.44	26.86		
120	**1868**	8.25	27.06		
121	**1869**	8.43	26.64	28.52	
122	**1870**	8.20	26.51	28.75	
123	**1871**	8.12	26.54	28.43	

b. Moving average: I choose 10 years is the averaging of moving as during around 10 years since the length of observation is pretty long (203 years)
c. Line charts:

The line chart shows the trend of average of temperature yearly. Y axis indicates the temperatures in C degree while the X axis shows the range of observation years in which the observation values in Perth city is available to pick up.


The trend of temperatue in some cities vs average global
Degree (C)
35
30
25
20
15
10
5
0
-5
-10
y = 0.0095x + 17.329
y = 0.008x + 7.8726
Years
1852
1857
1862
1867
1872
1877
1882
1887
1892
1897
1902
1907
1912
1917
1922
1927
1932
1937
1942
1947
1952
1957
1962
1967
1972
1977
1982
1987
1992
1997
2002
2007
2012
Global_moving_avg
Ho Chi Minh City_moving_avg
Khartoum_moving_avg
Perth_moving_avg
Ulaanbaatar_moving_avg
Linear (Ho Chi Minh City_moving_avg)
Linear (Perth_moving_avg)
Linear (Global_moving_avg)

4. Make observations:
    a. The trend of temperature is gradually increase over the time
    b. The trend of temperature in Perth city yearly is quite identical compare with the trend of temperature in global, indicated by two parallel trend line. So, from the observation's value of global data, it can be predicted the temperature of Perth city in that year.
    c. The range of average temperature in global varies from 7.203 to 9.594 degree compare with 16.21 to 19.28 in Perth
    d. From the years of 1970s, the trend of temperature in the world tends to crease faster compare with prior.