Assignment 1: Exploring Weather Trends - Project

1. Extract data

I pulled out all of the dataset by using SQL:

1. City\_data:

SELECT \* FROM city\_data

1. City\_list

SELECT \* FROM city\_list

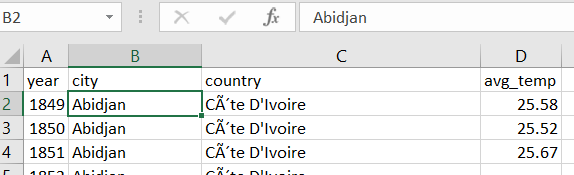
1. Global\_data

SELECT \* FROM global\_data

1. 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:

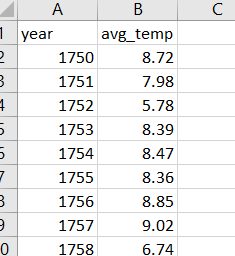
1. The result of city\_data: an excel file with 4 columns – year, city, country, and avg\_temp for every year in each city which the list is in city\_list file



1. City\_list file: An excel file with 2 columns – city and country correspond with this city

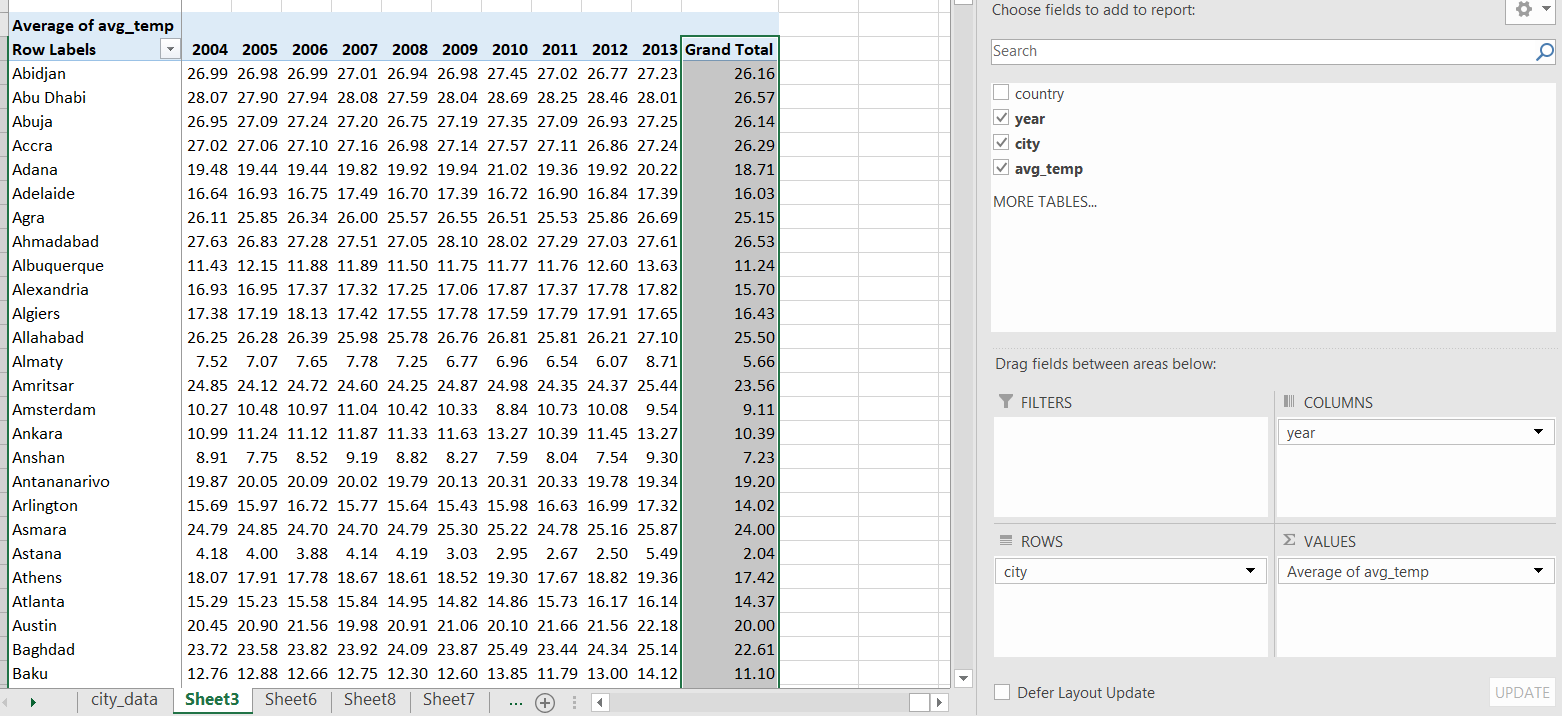


1. Global\_data: an excel file shows the global temperature’s observation each year, from 1750 o 2015



1. 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.



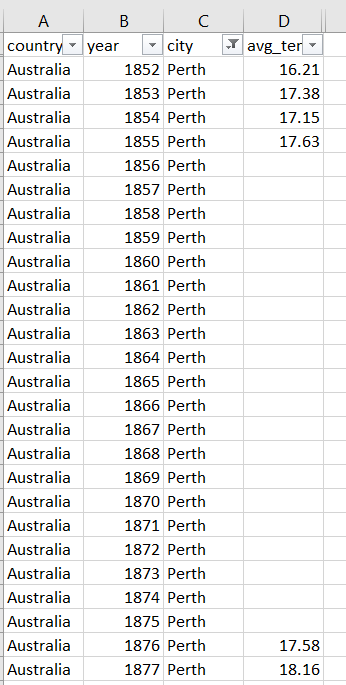
From the Grand Total column, which indicate the average of agv\_temp in each city, I extract the highest and lowest values and city corresponded with those values. So, the final selection to be appeared as list:

|  |  |  |
| --- | --- | --- |
| 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% |  |

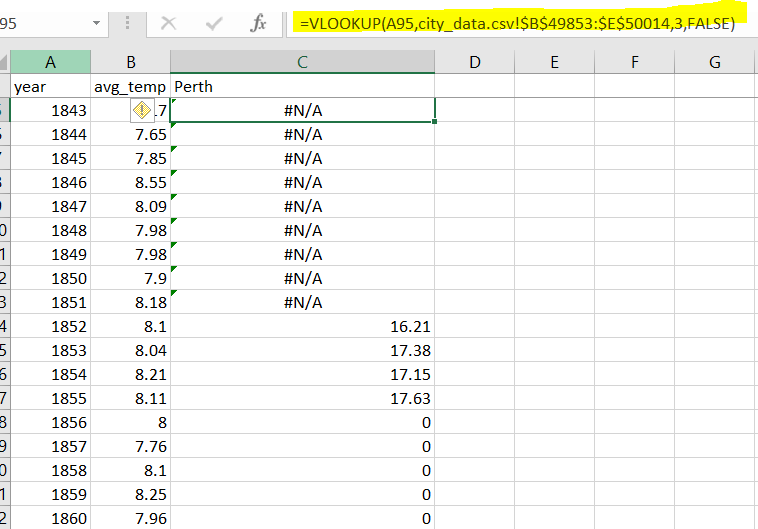
The table shows the name of 4 selected countries, indicate by total of non-missing values (the percentage of years the value is not missing compared with total observation’s years which taken in global data), total Counts of avg\_temp, and overral average of avg\_temp

1. Create the line chart
   1. Merge city\_data table with global\_data

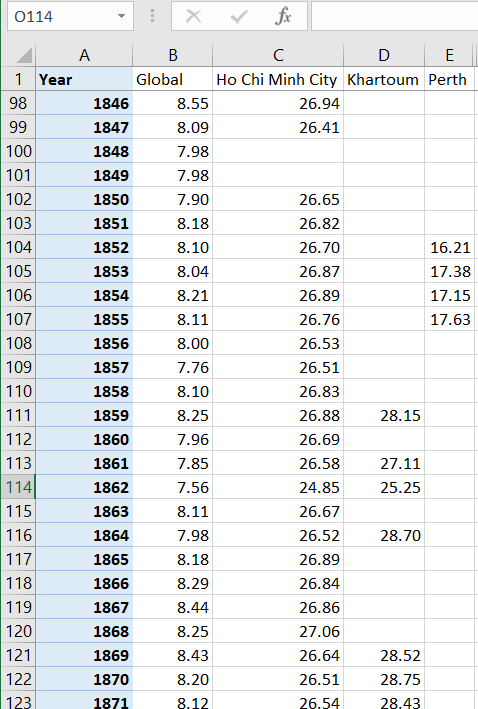
From city\_data excel -> Data -> Filter ->Select Perth from list of city name, we got the filter table of Perth’s temperature data over the observation years



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



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



* 1. 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)
  2. 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.

1. Make observations:
   1. The trend of temperature is gradually increase over the time
   2. 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.
   3. The range of average temperature in global varies from 7.203 to 9.594 degree compare with 16.21 to 19.28 in Perth
   4. From the years of 1970s, the trend of temperature in the world tends to crease faster compare with prior.