Udacity Data Analyst Nanodegree

First Project "Explore Weather Trends"

Outline of steps

Extract Data

1. First gather up the global temperature trends by running:
   select * from global_data;
   and export as global.CSV
2. Then gather up the temperature trends of my local area (Brussels, Belgium) by running:
   select * from city_data where country = 'Belgium' and city = 'Brussels'
   and export as brussels.CSV
3. Finally, just as an extra, check what is happening to the antipodes by gathering up the temperature trends of Sydney (Australia), so run:
   select * from city_data where country = 'Australia' and city = 'Sydney'
   and export as sidney.CSV

Open up the CSV

I don't have my usual laptop and so I have to use what comes with the spare one that my company provide me with, leading to few extra steps.

So I open each CSV in LibreOffice Calc then I consolidate all the data into one sheet that I export as an xlsx file. To do so I have to take care of the date range:
- global temperature trends cover a period from 1750 to 2015,
- the temperature trends of Brussels cover a period from 1743 to 2013,
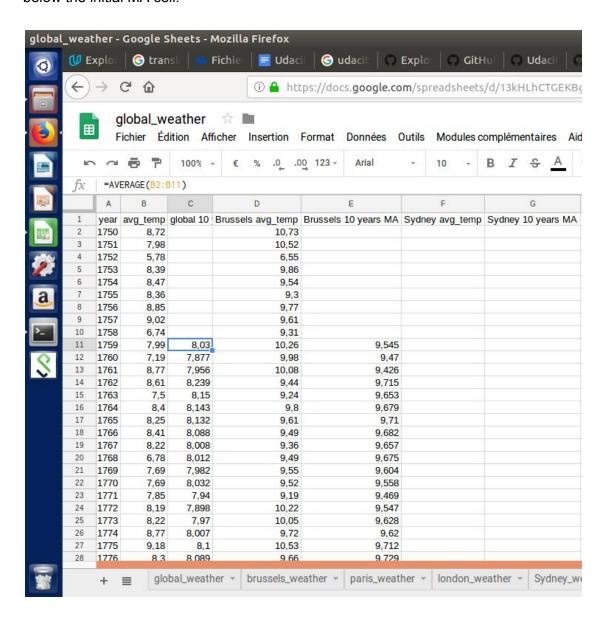- the temperature trends of Sydney cover a period from 1850 to 2013.

For consistency I keep the common date range for Global and Brussels. (I consider Sydney as an extra so I did not try to align data on the same time line).

Then I import that file into google docs, so that I can use tools I'm more use to.

Create the line chart

Create a 10 years moving average for Global, Brussels and also Sydney. To do so I just fill the following formula at the 11th line relatively to the region average temperature : "AVERAGE(X2:X11)" where X is the column of the region I'm working on (row numbering starts at 2 and ends at 11 since the first row is the header). Then I just select the bottom
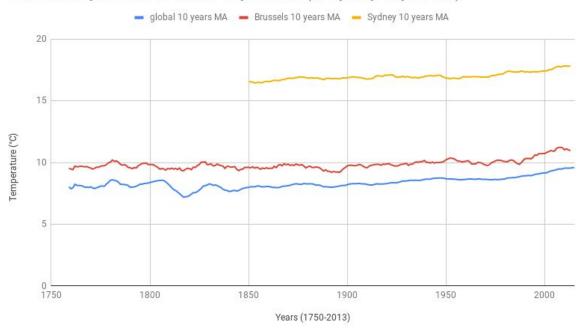
right edge of the cell and slide down up to the extent to apply the formula to all the rows below the initial MA cell.

Then just use the built-in graphic feature to plot the trends.

Line chart

Brussels 10 years MA vs Global 10 years MA (vs Sydney 10 year MA)

Observation

- Global and local temperature trends are consistent over the last few hundreds years : temperature is increasing, indicating that Earth is getting hotter.
- Brussels is slightly hotter that the global average, which is normal since it's located in a temperate zone. Sydney is by far hotter than Brussels and the global average. which is normal since it's next to the Tropic of Capricorne.
- Moving average is definitely more accurate to follow up trends, for example if I want to get an idea of how much Earth is getting hotter, I can subtract the temperatures from 1750 to 2013 and so end up with an amplitude of (9.61 - 8.72 =) **0.89°** when in fact, if I compare the MA from 1759 to 2013, I got an amplitude of (9.556 - 8.03 =) **1.526°** which is more realistic.
- In comparison Brussels temperature trends have an amplitude of (10.947 - 9.545 =) **1.402°**. Since I'm not a weather expert I don't know if a difference of **0.124°** is meaningful or not and so I won't make any conclusion on the relative global and brussels temperature. However by applying the same logic to Sydney, we end up with an amplitude of (17.844 - 16,538 =) **1.306°.** So I have the intuition that polar zones are getting hot faster than the other zones. To confirm this I should go through more data from polar to equatorial zone by following some gradient.

- Finally I notice that Brussels has not the same temperature drop pattern than the global trends between 1830 to 1844. That period corresponds to Belgium getting independence from Holland. More investigation is required to draw some conclusion about what is probably a difference in the industrial revolution of the country compare to the rest of the world.