Name: Guilherme Artur Lamb

In order to organize this work, it is divided in three topics in the following order:
1) Data gathering;
2) Data analysis;
3) Results.

---

1) Data gathering

These were the steps I followed in order to do the Weather Analysis:
The first step was to know which cities were available for Brazil. In order to do so, I made the following query:

```
SELECT city
FROM city_list
WHERE country = 'Brazil';
```

Where I filtered the results to narrow it down to just Brazil. After that, I noticed that Porto Alegre was available (which is the biggest city near me).

Then I made a query in the city_data table in order to get all data from Porto Alegre:

```
SELECT *
FROM city_data
WHERE city = 'Porto Alegre';
```

After that, I just made a query to obtain the global data:

```
SELECT *
FROM global_data;
```

---

2) Data analysis

I used Microsoft Excel to work the data. First, I created a spreadsheet for the data of Porto Alegre and one for the Global data. It were missing some temperature data for Porto Alegre, so I excluded the years that the data were missing from this study (so it do not prejudicate my average). As we can see in figure 01.

	A	B	C	D
1	year	city	country	avg_temp
2	1832	Porto Alegre	Brazil	17.37
3	1833	Porto Alegre	Brazil	18.14
4	1834	Porto Alegre	Brazil	17.79
5	1835	Porto Alegre	Brazil	17.05
6	1836	Porto Alegre	Brazil	17.38
7	1837	Porto Alegre	Brazil	16.98
8	1838	Porto Alegre	Brazil	17.21
9	1839	Porto Alegre	Brazil	17.19
10	1840	Porto Alegre	Brazil	17.62
11	1841	Porto Alegre	Brazil	17.40
12	1842	Porto Alegre	Brazil	17.81
13	1843	Porto Alegre	Brazil	17.93
14	1844	Porto Alegre	Brazil	
15	1845	Porto Alegre	Brazil	
16	1846	Porto Alegre	Brazil	
17	1847	Porto Alegre	Brazil	
18	1848	Porto Alegre	Brazil	
19	1849	Porto Alegre	Brazil	
20	1850	Porto Alegre	Brazil	
21	1851	Porto Alegre	Brazil	17.90
22	1852	Porto Alegre	Brazil	17.93
23	1853	Porto Alegre	Brazil	17.92
24	1854	Porto Alegre	Brazil	18.06
25	1855	Porto Alegre	Brazil	17.94
26	1856	Porto Alegre	Brazil	17.63
27	1857	Porto Alegre	Brazil	18.21

Figure 1 - Missing data

After that, I've calculated the moving average for two and four years. And to see which one was best, I used the method of the mean squared error (MSE). (* Média = Average)

F3 =MÉDIA(D2:D3)

	A	B	C	D	E	F	G
1	year	city	country	avg_temp		2 years Moving Average	4 years Moving Average
2	1832	Porto Alegre	Brazil	17,37			
3	1833	Porto Alegre	Brazil	18,14		17,755	
4	1834	Porto Alegre	Brazil	17,79		17,965	
5	1835	Porto Alegre	Brazil	17,05		17,42	17,5875
6	1836	Porto Alegre	Brazil	17,38		17,215	17,59
7	1837	Porto Alegre	Brazil	16,98		17,18	17,3
8	1838	Porto Alegre	Brazil	17,21		17,095	17,155
9	1839	Porto Alegre	Brazil	17,19		17,2	17,19
10	1840	Porto Alegre	Brazil	17,62		17,405	17,25
11	1841	Porto Alegre	Brazil	17,4		17,51	17,355
12	1842	Porto Alegre	Brazil	17,81		17,605	17,505
13	1843	Porto Alegre	Brazil	17,93		17,87	17,69
14	1851	Porto Alegre	Brazil	17,9		17,915	17,76
15	1852	Porto Alegre	Brazil	17,93		17,915	17,8925
16	1853	Porto Alegre	Brazil	17,92		17,925	17,92
17	1854	Porto Alegre	Brazil	18,06		17,99	17,9525
18	1855	Porto Alegre	Brazil	17,94		18	17,9625

Figure 2 – 2 years moving average

G5 =MÉDIA(D2:D5)

	A	B	C	D	E	F	G
1	year	city	country	avg_temp		2 years Moving Average	4 years Moving Average
2	1832	Porto Alegre	Brazil	17,37			
3	1833	Porto Alegre	Brazil	18,14		17,755	
4	1834	Porto Alegre	Brazil	17,79		17,965	
5	1835	Porto Alegre	Brazil	17,05		17,42	17,5875
6	1836	Porto Alegre	Brazil	17,38		17,215	17,59
7	1837	Porto Alegre	Brazil	16,98		17,18	17,3
8	1838	Porto Alegre	Brazil	17,21		17,095	17,155
9	1839	Porto Alegre	Brazil	17,19		17,2	17,19
10	1840	Porto Alegre	Brazil	17,62		17,405	17,25
11	1841	Porto Alegre	Brazil	17,4		17,51	17,355
12	1842	Porto Alegre	Brazil	17,81		17,605	17,505
13	1843	Porto Alegre	Brazil	17,93		17,87	17,69
14	1851	Porto Alegre	Brazil	17,9		17,915	17,76
15	1852	Porto Alegre	Brazil	17,93		17,915	17,8925
16	1853	Porto Alegre	Brazil	17,92		17,925	17,92
17	1854	Porto Alegre	Brazil	18,06		17,99	17,9525
18	1855	Porto Alegre	Brazil	17,94		18	17,9625

Figure 3 - 4 years moving average

In figure 4, we can see that the error was least for the 2 years moving average than the 4 years. So for this work I selected the 2 years moving average.

2004	Porto Alegre	Brazil	18,75		18,91	19,26
2005	Porto Alegre	Brazil	19,42		19,085	19,1825
2006	Porto Alegre	Brazil	19,35		19,385	19,1475
2007	Porto Alegre	Brazil	19,05		19,2	19,1425
2008	Porto Alegre	Brazil	18,8		18,925	19,155
2009	Porto Alegre	Brazil	18,84		18,82	19,01
2010	Porto Alegre	Brazil	19,01		18,925	18,925
2011	Porto Alegre	Brazil	18,82		18,915	18,8675
2012	Porto Alegre	Brazil	19,72		19,27	19,0975
2013	Porto Alegre	Brazil	17,84		18,78	18,8475
					2 years Moving Average	4 years Moving Average
				Mean Squared Error	0,062108477	0,111325291

Figure 4 - Mean squared error

For the calculation of the MSE I've used the following formula:

$$\frac{1}{N}\sum_{i=1}^{N}(Y_i - Y_p)^2$$

Where:

N – is the sample size;

Y_i – is the vector of observed values of the variable being predicted;

Y_p – is the predicted values

The process for analysing the global temperature data was the same, except that I only have calculated the 2 years moving average, in order to be equal to the Porto Alegre process.

Then I've ploted the line chart for the moving average for the temperature of Porto Alegre and Global. I also used a tendency line to see the increase in the temperature. In this way is easier to detect any trend and a line chart is less polluted than a bar chart and suits us better.

---

3) Results

First shown here is the graphic for Porto Alegre temperature (figure 5 below). We can see that the temperature was around 17,5 ℃ and 18 ℃ during most part of the first century of data recording (years 1833 until 1926). After that we can clearly see a rapid increase of temperature, with almost all averages above 18 ℃ and after year 2000 the average was almost all the time above 19 ℃.

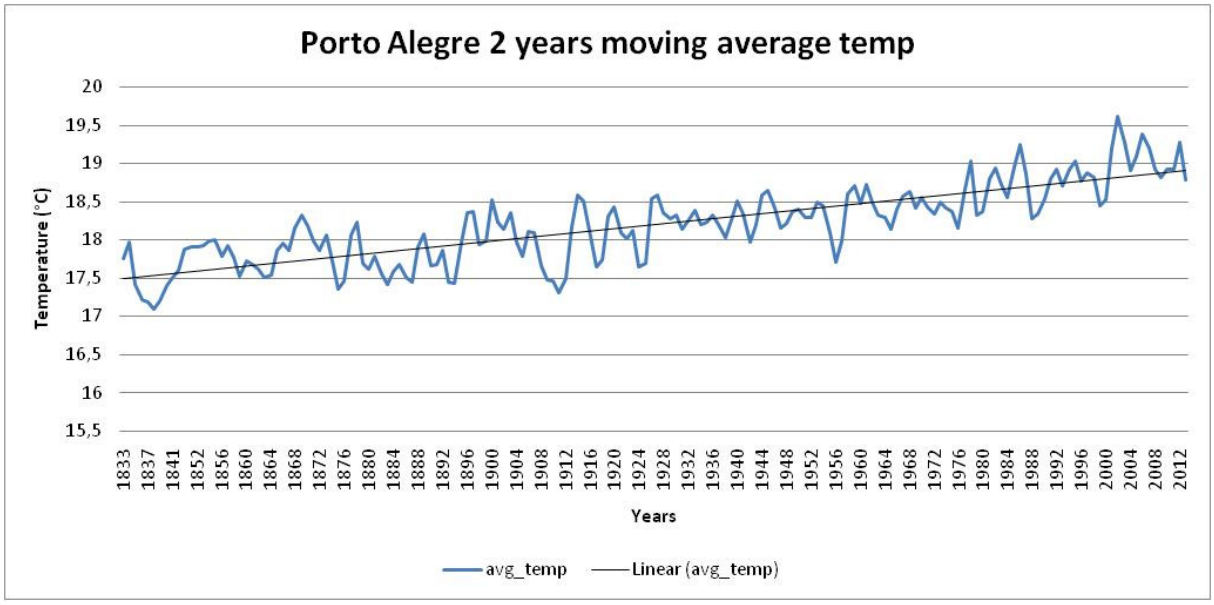


Figure 5 - Average temperature in Porto Alegre

Another factor is the deforestation of green areas for housing construction, currently Porto Alegre is the 12th most populous city in Brazil, with about 1.5 million inhabitants.

Next is presented the chart for the average global temperature.

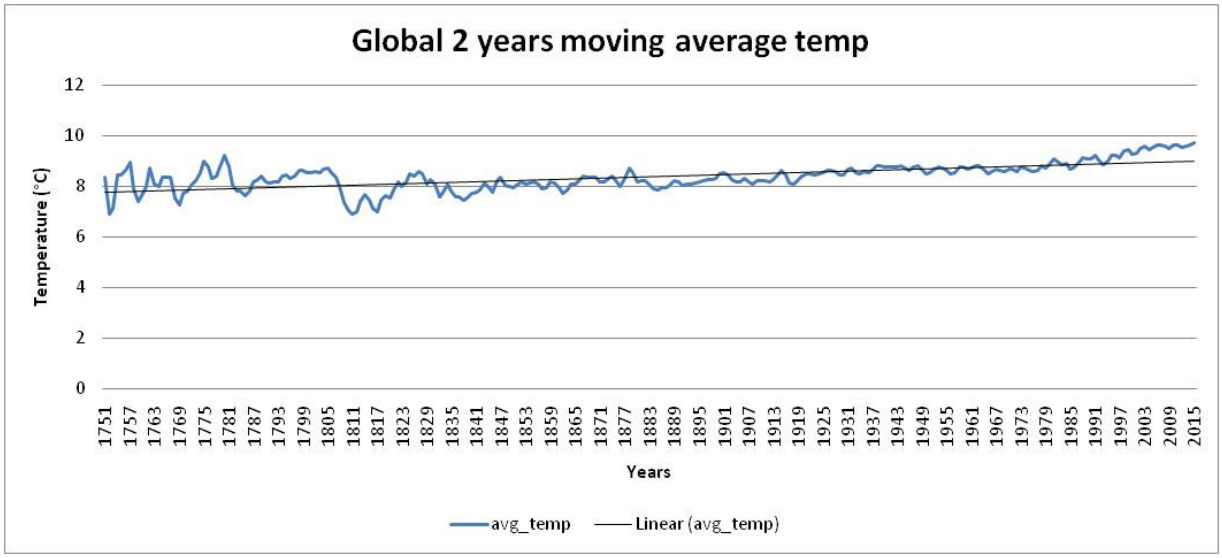


Figure 6 - Global avarage temperature

As we can see, this chart also presented a increasing tendency. During basically the first 60 years of measuring we have a lot of fluctuation, even with the method of the moving average. One explanation for this could be the methods that were used do collect and save this data, we cannot know if they were precise or not. After year 1820 we have a more smooth line in the chart. During the years 1800 the increase of the temperature was very subtle, only after around 1920 we can notice a more rapid increase, probably because of industrialization and the agglomeration of people in large urban centers. Around the year of 1985, the increase in temperature began to accelerate significantly, increasing by about 1 ℃ in thirty years, being close to 10 ℃.

The figure 7 below shows us the combined average temperature for Porto Alegre and Global.

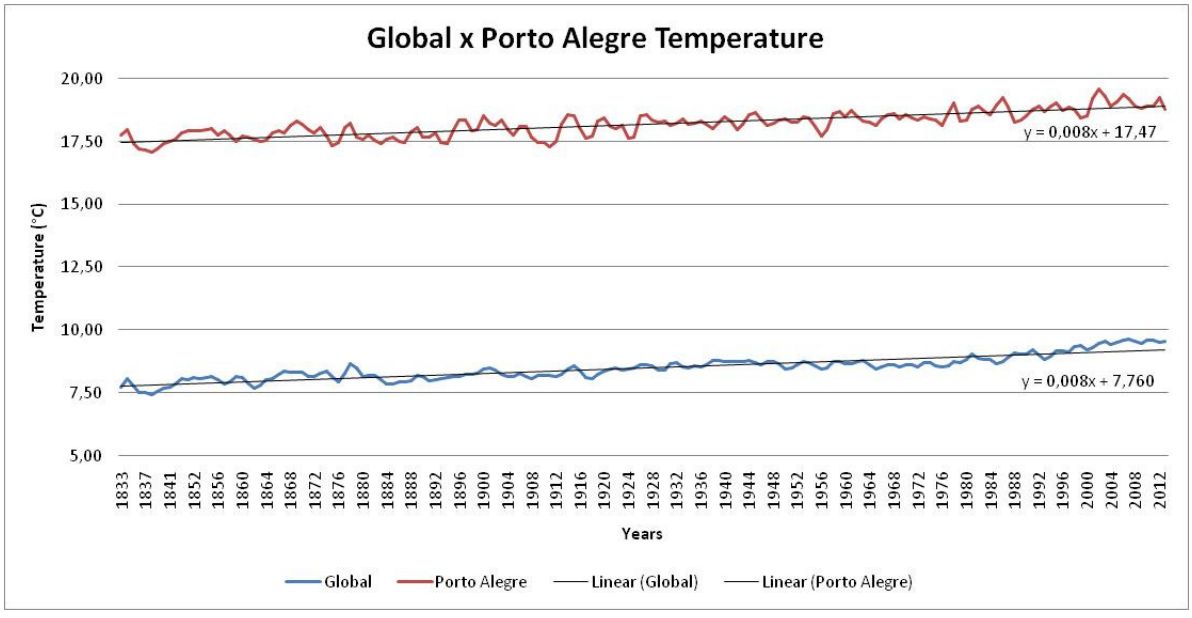
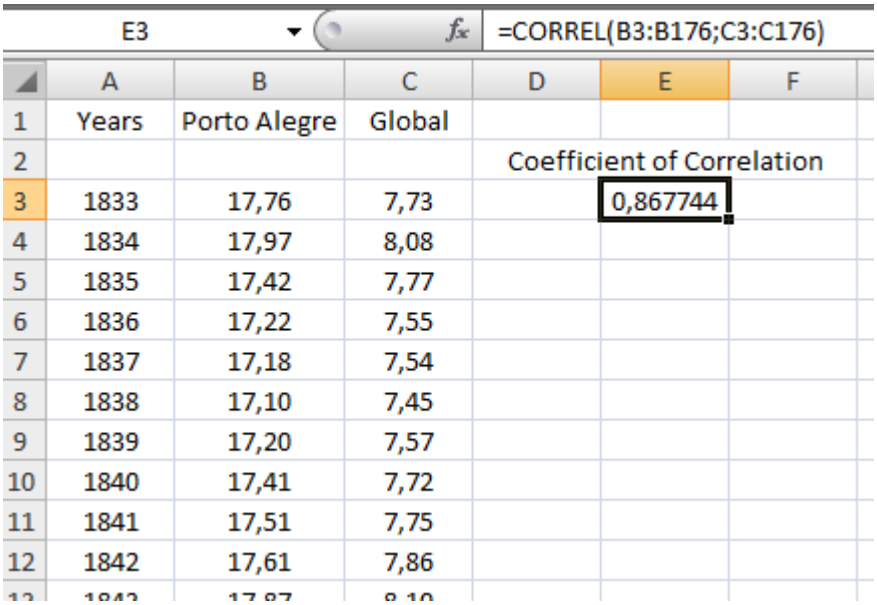


Figure 7 - Average Global temperature and for Porto Alegre

To do this graphic, I have selected the years of the Global average temperature that corresponds to the ones of Porto Alegre. So the years are the same and the size of the sample is equal for both.

As we can see, the average temperature in Porto Alegre is about 10 ℃ higher than the Global average and this difference seems to be consistent over the years. Both have presented the same increase tendency. Looking at the graphic, we can observe that the equation for the linear tendency line gives the same angular coefficient for both cases (x=0,008). This trend seems to be constant, but it seems to have accelerated from the 70s and 80s for both cases.

Adding to this, I have calculated the coefficient of the correlation between the data, thru the Excel function CORREL, and the result showed that they have a strong correlation (~0.87). With this, we could assume that a increase in the Global temperature would also increase the temperature in Porto Alegre.



E3 | =CORREL(B3:B176;C3:C176)

	A	B	C	D	E	F
1	Years	Porto Alegre	Global			
2				Coefficient of Correlation		
3	1833	17,76	7,73		0,867744	
4	1834	17,97	8,08			
5	1835	17,42	7,77			
6	1836	17,22	7,55			
7	1837	17,18	7,54			
8	1838	17,10	7,45			
9	1839	17,20	7,57			
10	1840	17,41	7,72			
11	1841	17,51	7,75			
12	1842	17,61	7,86			

Figure 8 - Correlation analysis