## Project: Explore Weather Trends

Nanodegree Udacity Program: Data Analyst

Student: Abderrahman BENYAHYA

Morocco, Rabat

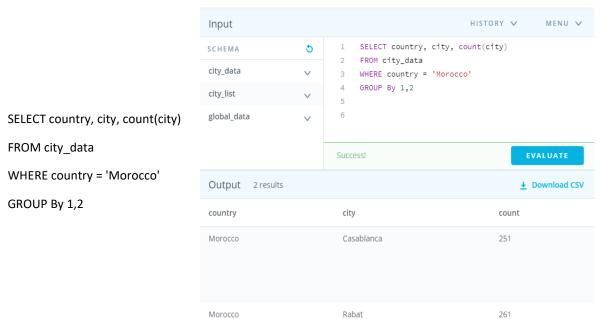
17 May 2019

1. An outline of steps taken to prepare the data to be visualized in the chart, such as: What tools did you use for each step? (Python, SQL, Excel, etc)

During the project, I had first to extract the data from the Database in the classroom using the SQL commands. The first difficulty was to identify which city should be considered as mine.

I entered first the SQL command to verify if there are and how many cities from Morocco using the

code:



It returns the results that there are two cities Casablanca and Rabat. So I write the second code:

SELECT year, city, avg\_temp

FROM city\_data

WHERE city = 'Rabat'

I uploaded the result in csv and created the chart using Excel 2019 professional but also Open Office to have the csv converted properly to odt and xlx.

## Project: Explore Weather Trends

I did the same with the global data.

SELECT year, avg\_temp

FROM global\_data

Then I created the chart by calculating the average for ten years as indicated in the classroom. I used the Excel function =averagea() of the first ten years of the global and Rabat's figure and pull that it calculates automatically for the rest.

D11 • : × ✓ fx =AVERAGEA(B2:B11)										
4	Α	В	С	D	Е	F	G	Н	1	J
1	Year	Global	Rabat	AVG global	AVG Rabat	Ecart Global	Ecart Rabat			
2	1750	8,72				0,35052632				
3	1751	7,98				-0,38947368				
4	1752	5,78				-2,58947368			Global AVG:	8,36947368
5	1753	8,39	17,31			0,02052632	-0,0120307		Rabat AVG:	17,3220307
6	1754	8,47	17,47			0,10052632	0,1479693			
7	1755	8,36	17,16			-0,00947368	-0,1620307			
8	1756	8,85	17,53			0,48052632	0,2079693			
9	1757	9,02	17,42			0,65052632	0,0979693			
10	1758	6,74	15,89			-1,62947368	-1,4320307			
11	1759	7,99	17,01	8,03		-0,37947368	-0,3120307			
12	1760	7,19	16,38	7,877		-1,17947368	-0,9420307			
13	1761	8,77	17,41	7,956		0,40052632	0,0879693			
14	1762	8,61	17,41	8,239		0,24052632	0,0879693			
15	1763	7,5	16,55	8,15	17,099	-0,86947368	-0,7720307			
16	1764	8,4	17,41	8,143	17,023	0,03052632	0,0879693			
17	1765	8,25	17,43	8,132	17,017	-0,11947368	0,1079693			

But I found out that it would be more efficient to have the difference between each temperature and the average of temperature for the entire time. I calculated the average for the all period (using the function =averagea(). It gave 8.36947368 for Global and 17.3220307 for Rabat. Then I calculate the difference between the temperature of each year and the average (the value).

The most effective visualization in line diagram as it should reflect the fluctuation by year. The color of the two elements (Rabat and Global) should be distinctive.

I used R and R Studio to import the csv then to calculate the Pearson correlation between the two data (Rabat and Global)

>smp = read.csv2(file = "compare Rabat Global.csv", dec=',')

>cor.test(smp\$AVG.global,smp\$AVG.Rabat)

## Project : Explore Weather Trends

2. Line chart with local and global temperature trends

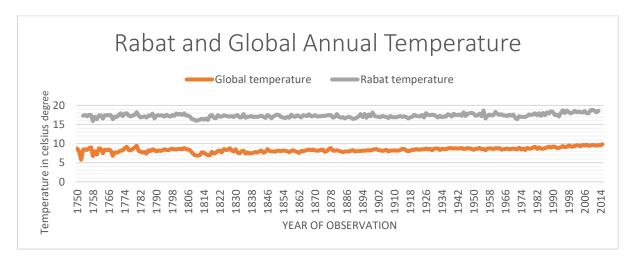


Fig1. Annual temperature for Rabat and Global

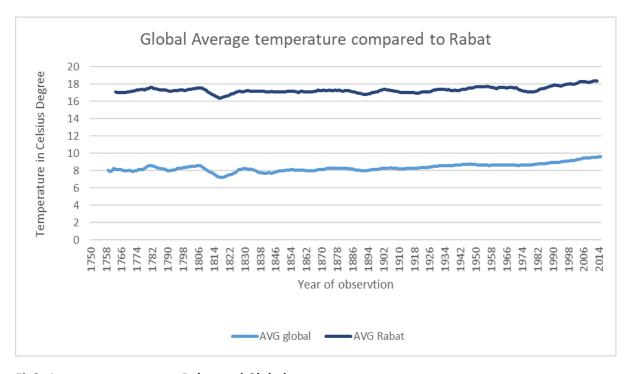


Fig2. Average temperature Rabat and Global ten years.

## Project : Explore Weather Trends

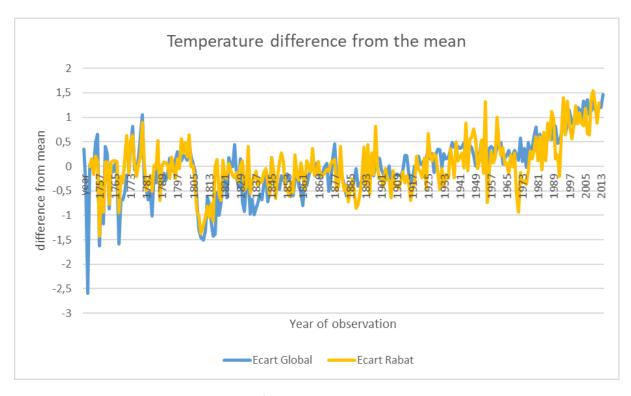


Fig3. Distance to the mean temperature for Rabat and Global

- 3. At least four observations about the similarities and/or differences in the trends
- The two first line charts are very similar even if the average tends to erase the discontinuity
- We can see that Rabat is always warmer than the global average temperature
- The two lines are following almost the same trend both in up and down. The person's product correlation calculated with R is 0.8971999, very strong correlation to conclude that Rabat is following the global trends.
- Using the two first chart does not give a really good picture of the climate change over the years. This is why it is preferred to calculate the average of temperature for the all period and see for each year the distance to mean.
- The last chart shows clearly that there is a correlation between the two line (Rabat and Global). The global warming is very clear and shown by the increase of temperate since almost 1909. It seems that the distance was never so high in history as it is since 1989.