Explore Weather Trends

Summary

In this project, I analyzed local and global temperature data and compared the temperature trends from my hometown of Houston to overall global temperature trends.

Explore Part 1

Extract

From the Udacity PostgreSQL workspace, I wrote the following query to extract global temperature data from my hometown of Houston.

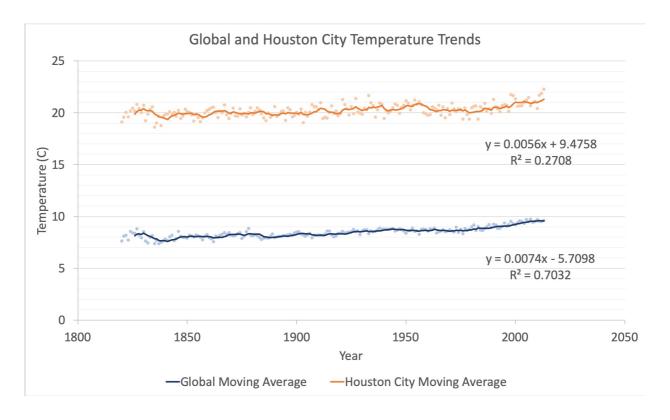
```
SELECT
    gd.year AS year,
    gd.avg_temp AS global_temp,
    cd.city AS city,
    cd.country AS country,
    cd.avg_temp AS city_temp
FROM
    global_data gd
    JOIN city_data cd ON gd.year = cd.year
    JOIN city_list cl ON cl.city = cd.city
WHERE
    cl.city = 'Houston';
```

Visualize

I used Excel to help visualize the data set. I calculated the moving average by looking at the previous 7 years using the following formula:

	Α	В	С	D		
1	year	global_temp	global_mov_avg	city		
2	1820	7.62		Houston		
3	1821	8.09		Houston		
4	1822	8.19		Houston		
5	1823	7.72		Houston		
6	1824	8.55		Houston		
7	1825	8.39		Houston		
8	1826	8.36	=AVERAGE(B2:B8)	Houston		
9	1827	8.81	8.301428571	Houston		

I displayed the yearly data as points in the following graph. Afterwards, I created a line graph with the moving average data. Excel also has a built in function to create a trend-line for the data as well as calculating the coefficient of determination or R² value.



From the above chart, I was able to make the following inferences:

- 1. The average Houston temperature is roughly 10 °C hotter than the global temperature.
- 2. For the years 1820-2013, the average global temperature is increasing roughly 0.74 °C every hundred years.
- 3. For the years 1820-2013, the average Houston temperature is increasing roughly 0.56 $^{\circ}$ C every hundred years.

Explore Part 2

Extract

Upon further inspection, I noticed that certain cities had data starting in the 1700's. Just for fun, I decided to update my SQL query to dig into the data a bit more and to also include all the cities I have visited in the past.

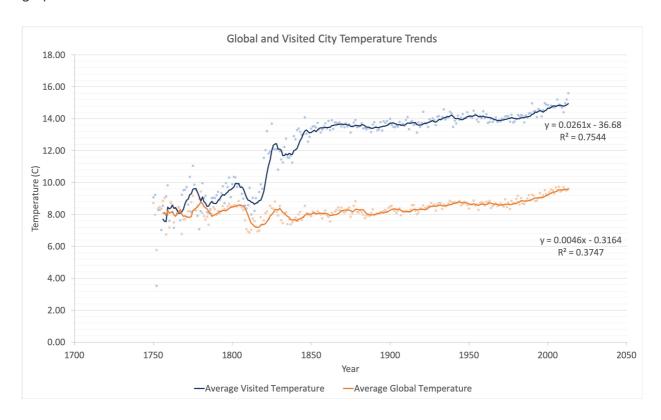
From this new data subset, I was able to create the following Pivot Table with cities as row, years as columns, and average temperature as values.

Once we reach the mid 1800's, almost all cities have average temperature data.

		Column Labels -T													
low Labels	-7	1750	1751	1752	1753	1754	1795	1796		1758	1759	1760	1761	1762	1263
Amsterdam		30.04	9.63	5.97	9.08	8.72	8.55	9.17		8.79	9.64	9.14	9.41	8.58	8.44
Arlington		12.51	13.26	5.45	11.91	12.22	9.15	12.36	11.36	20.48	11.43	30.1	12.55	11.92	9.7
Austin															
Bangkok															
ten		7.61	7.73	2.86	6.83	6.55	6.21	6.6	6.3	5.92	7.06	7.27	7.08	6.64	6.43
Boston		7.88	8.6	0.36	2.36	2.75	4.28	7.36	6.65	6.09	6.8	5.53	8.05	2.42	4.99
Budspent		30.55	10.89	4.74	9.66	9.55	9.35	50.34	9:92	8.71	9.73	9.49	10.13	9.72	9.22
Ovcago		30.49	11.19	4.5	10:04	30.64	7.41	11.15	9.5	8.61	9.59	8.05	10.4	9.82	7.72
Dallas															
Devreed															
Fort Worth															
Helyarkii		5.14	4.68	-0.29	4.14	4.15	4.05	4.47	4.75	2.66	4.08	2.79	4.55	4.22	3.35
HI ON MAN CRY													-		
Houston															
London		10.25	9.99	6.54	9.42	9.2	8.95	9.42	9.34	8.85	9.8	9.36	9.69	8.99	8.75
Long Beach			****	-	-										
Los Angeles															
Miami										23.05	22.56	15.14			
Milan		7.51	7.77	2.65	6.83	6.53	6.27	6.65	6.3	5.97	7.14	7.15	6.99	6.69	6.48
Montreal		5	5.62	-3.22	4.44	4.54	1.16	4.8	3.7	3.14	3.92	2.36	4.58	4.66	1.62
New Orleans			2.44	-0.00	4.44	4.54	2.10		2.7	20.87	19.79	9.34		4.00	1.64
New York		10.07	10.79	2.81	9.52	9.88	6.60	9.94	8.89	8.15	9.00	7.73	10.18	9.55	7.23
Oldahoma City		80.00	90.79	2.00	9.54	7.66	0.00	9,54	***	8.15	9.00	7.79	10.00	9.00	7.45
		3.57	3.06	-2.24	2.15	2.18	1.6	2.30	2.52	1.45	2.64	1.75	2.97	1.9	1.91
Oslo Paris				6.97	30.4	10.15	9.84	10.2		9.62	20.66	10.5			
		11.18	11.15										10.55	10.06	9.83
Philadelphia		12.36	13.05	5.64	11.85	12.29	9.12	12.41	11.26	10.43	11.35	10.12	12.47	11.8	9.63
Phoenix															
Rome		12.53	12.99	7.68	11.96	11.79	11.57	12.13	11.8	11.02	12.12	11.94	13.21	11.99	11.54
Sacramento															
San Antonio															
San Diego															
San Francisco															
Seuttle															
Secoul															
Shanghai															
Singapore															
Stockholm		7.35	6.46	3.06	6.18	6.16	5.82	6.3	6.51	5.21	6.58	5.62	6.88	6.23	5.61
Talpel															
Tokyo															
Tuba										26.67	34.63	-0.55			
Vienna		8.86	9.03	3.42	7.91	7.73	7.53	8.51	8.14	7.13	8.09	7.99	8.46	7.89	7.58

Visualize

I took the average of all city temperatures for the year and used Excel to visualize the following graph:



My initial reaction to the spike around the year 1820 was that maybe it was caused by the industrial revolution.

However, on closer inspection, it appears that a lot of the cities in my home state of Texas (it's hot here!) began reporting their average temperature in that year.

Interpret

From the above chart, I was able to make the following inferences:

- 1. Temperature data prior to 1800 had a relatively large standard deviation compared to the data after 1850.
- 2. In the 21st century, I have personally visited many cities that are on average 6 °C warmer than the global average.
- 3. For the years 1750-2013, the average global temperature is increasing roughly 0.46 °C every hundred years. Compared to out previous visualization, it appears that the global temperature rate of change is increasing.

License

This project is licensed under the MIT License - see the LICENSE.md file for details.

Additional Notes

Since there are some duplicate city names, I manually deleted the data for London, Canada and Los Angeles, Chile since I have never been to these cities.