

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.

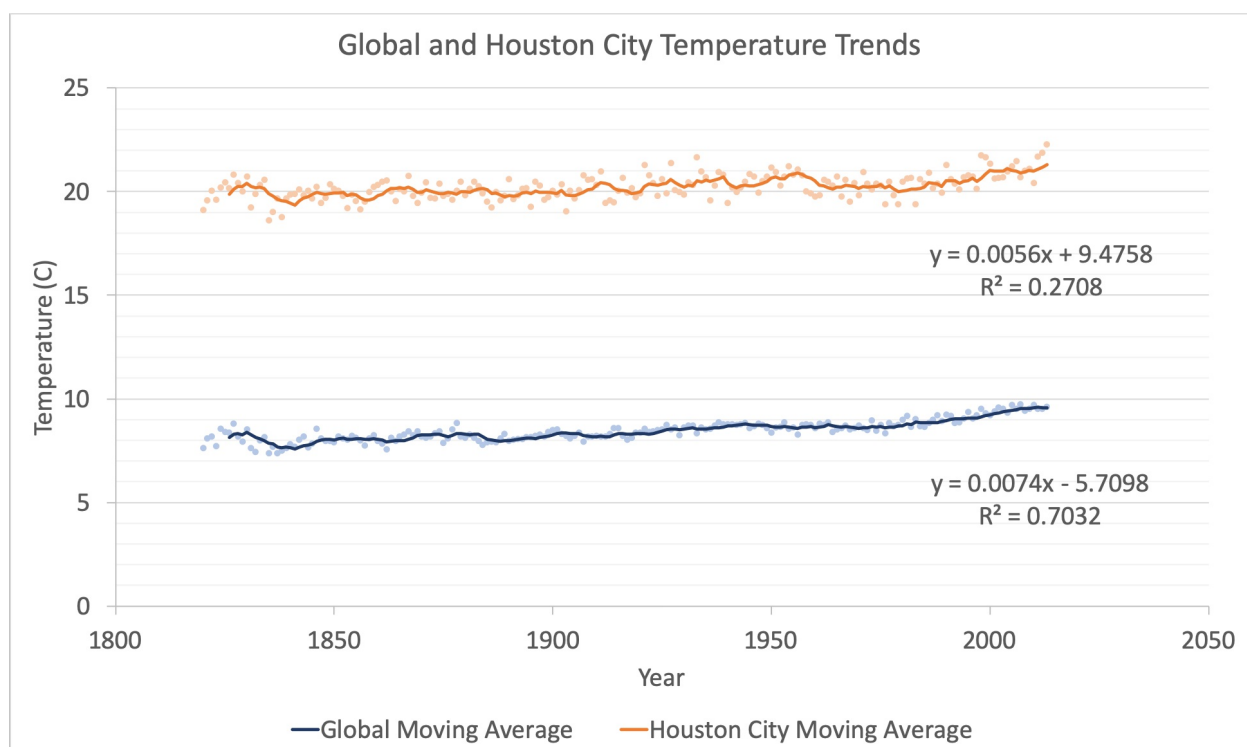
```
SELECT
    gd.year AS year,
    gd.avg_temp AS global_temp,
    cd.city AS city,
    cd.country AS country,
    cd.avg_temp AS city_temp
FROM
    global_data gd
    JOIN city_data cd ON gd.year = cd.year
    JOIN city_list cl ON cl.city = cd.city
WHERE
    cl.city = 'Houston';
```

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:

	A	B	C	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 R^2 value.



Interpret

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 °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.

```
WHERE
    cl.city IN ('Amsterdam', 'Arlington', 'Austin', 'Bangkok', 'Bern', 'Boston',
               'Budapest', 'Chicago', 'Dallas', 'Denver', 'Fort Worth',
               'Helsinki', 'Ho Chi Minh City', 'Houston', 'London',
               'Long Beach', 'Los Angeles', 'Miami', 'Milan', 'Montreal',
               'New Orleans', 'New York', 'Oklahoma City', 'Oslo', 'Paris',
               'Philadelphia', 'Phoenix', 'Rome', 'Sacramento', 'San Antonio',
               'San Diego', 'San Francisco', 'Seattle', 'Seoul', 'Shanghai',
               'Singapore', 'Stockholm', 'Taipei', 'Tokyo', 'Tulsa', 'Vienna');
```

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.

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 our 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.