EXPLORE WEATHER TRENDS

Udacity Data Analyst Nanodegree

Elvis McNeely

June 12, 2019

Term 01 - Project 01

Explore Weather Trends

Overview

In this report I extract, analyze and visualize local temperature trends (Keller, TX USA) compared to global temperature trends based on the defined set of goals. The data utilized for this report was provided from Udacity's learning platform.

GOALS

- Extract proper local and global data from provided sources, without errors
- Calculate ten year moving averages for the temperature data
- Create a line chart with moving averages local and global temperature data

TOOLS

- SQL, via the inline SQL editor connected to Udacity's data sources
- Excel, to calculate moving averages and plot line graphs

DATA SCHEMAS

- city_list: with column names city and country
- city_data: with column names year, city, country and avg_temp
- global_data: with column names year and avg_temp

STEP 1 - COLLECT THE DATA

1. Evaluate the SQL, extract the list of cities in the United States, to find my city or a city close to my physical location (Keller, TX USA). Based on the cities available in the data set, the closet city to Keller is Fort Worth, TX.

   ```
   SELECT * FROM city_list
   WHERE country = 'United States'
   ```

2. Evaluate the SQL, extract global and city records, download and save the results as results.csv. Both global_data and city_data tables have the year column in common. After evaluating the schemas, I decided to combine data from both tables. This scenario is valid because there are no duplicate records per year in either data set. If there were duplicate records, we might not be able to use this approach. ***Note: Forth Worth data is only available between years 1820 to 2013. We want to extract only the common years from both city and global data sets.***

```
SELECT
    cd.year,
    cd.avg_temp AS local_temp,
    gd.avg_temp AS global_temp
FROM city_data AS cd
LEFT JOIN global_data AS gd ON gd.year = cd.year
WHERE
    cd.country = 'United States' AND
    cd.city = 'Fort Worth' AND
    cd.year >= '1820' AND
    cd.year <= '2013'
ORDER BY cd.year
```

STEP 2 - EXCEL SPREADSHEET

1. Open the results.csv file and convert it to xlsx format. The spreadsheet will have four columns: local_year, local_temp, global_year, global_temp. The end result is:

	A	B	C	D
1	local_year	local_temp	global_year	global_temp
2	1820	16.88	1820	7.62
3	1821	17.33	1821	8.09
4	1822	17.87	1822	8.19
5	1823	17.46	1823	7.72
6	1824	17.9	1824	8.55
7	1825	18.38	1825	8.39
8	1826	17.93	1826	8.36
9	1827	18.62	1827	8.81
10	1828	18.26	1828	8.17
11	1829	17.89	1829	7.94
12	1830	18.68	1830	8.52
13	1831	16.98	1831	7.64
14	1832	17.81	1832	7.45
15	1833	18.29	1833	8.01

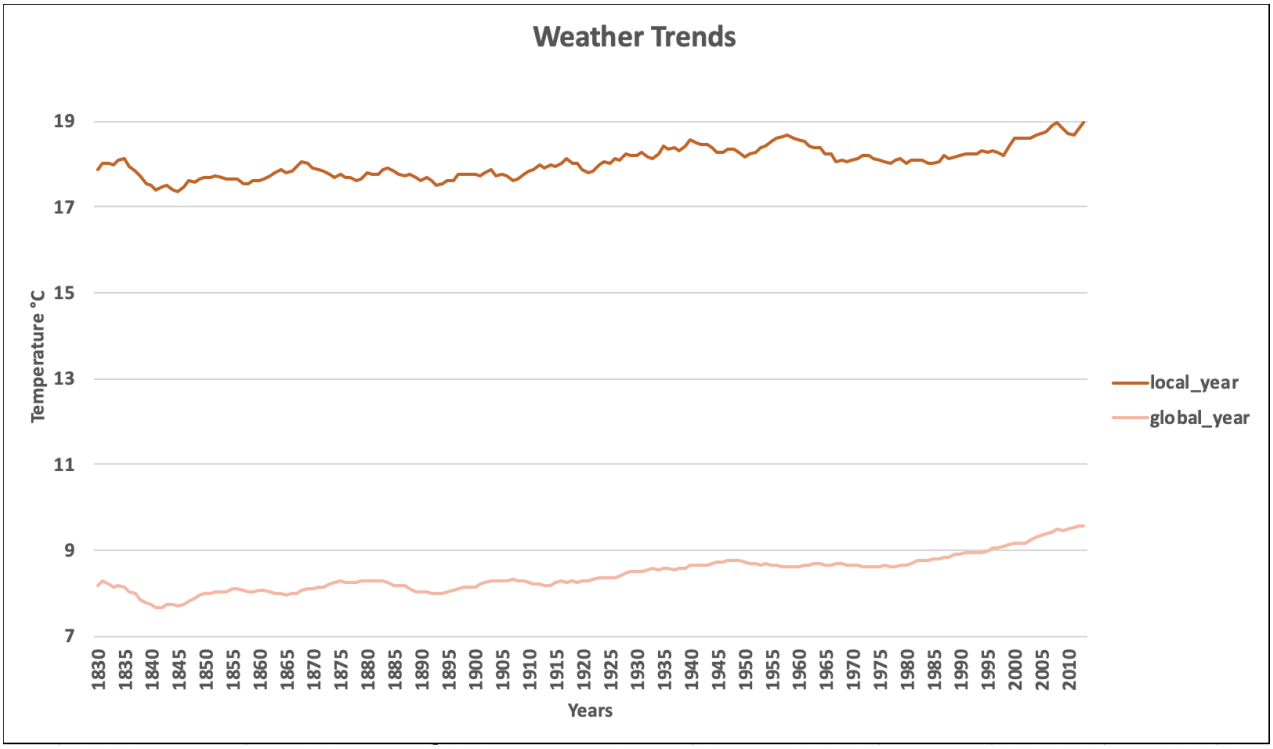
2. Add two new columns to hold the moving average for local_temp and global_temp. Then calculate a 10 year moving average for both columns:

   =AVERAGE(C2:C11)

   The above calculation above is for B12 cell in the **below** diagram. The end result is:

	A	B	C	D	E	F
1	local_year	local_temp_MA_10yr	local_temp	global_year	global_temp_MA_10yr	global_temp
2	1820		16.88	1820		7.62
3	1821		17.33	1821		8.09
4	1822		17.87	1822		8.19
5	1823		17.46	1823		7.72
6	1824		17.9	1824		8.55
7	1825		18.38	1825		8.39
8	1826		17.93	1826		8.36
9	1827		18.62	1827		8.81
10	1828		18.26	1828		8.17
11	1829		17.89	1829		7.94
12	1830	17.852	18.68	1830	8.184	8.52
13	1831	18.032	16.98	1831	8.274	7.64
14	1832	17.997	17.81	1832	8.229	7.45
15	1833	17.991	18.29	1833	8.155	8.01
16	1834	18.074	18.5	1834	8.184	8.15
17	1835	18.134	16.58	1835	8.144	7.39
18	1836	17.954	16.67	1836	8.044	7.7
19	1837	17.828	17.47	1837	7.978	7.38
20	1838	17.713	16.54	1838	7.835	7.51

3. Next, we need to graph the local_temp and global_temp using the ten-year moving averages.



OBSERVATIONS

- Local Temperatures

  Summary: Forth Worth has increased in temperature by 1.13 °C over the course of 184 years, with an average temperature **increase** of 0.006% per year.

  - Min: 17.352 °C (year of 1845)
  - Max: 18.984 °C (year of 2013)
  - Range: 1.632 °C
  - Mean: 18.037 °C

- Global Temperatures

  Summary: The global temperature has increased by 1.36 °C over the course of 184 years, with an average temperature **increase** of 0.007% per year.

  - Min: 7.69 °C (year of 1841)
  - Max: 9.56 °C (year of 2012)
  - Range: 1.87 °C
  - Mean: 8.443 °C

- Correlation Coefficient: 0.879

  Indicate both local and global temperatures have a close linear relationship. *(between local and global 10 year MA temperatures)*

	A	B	C	D	E	F
1	local_year	local_temp_MA_10yr	local_temp	global_year	global_temp_MA_10yr	global_temp
182	2000	18.595	18.9	2000	9.156	9.2
183	2001	18.588	18.55	2001	9.153	9.41
184	2002	18.596	18.19	2002	9.176	9.57
185	2003	18.596	18.55	2003	9.249	9.53
186	2004	18.669	18.58	2004	9.315	9.32
187	2005	18.696	19.13	2005	9.343	9.7
188	2006	18.764	19.85	2006	9.378	9.53
189	2007	18.906	18.42	2007	9.427	9.73
190	2008	18.963	18.52	2008	9.48	9.43
191	2009	18.83	18.42	2009	9.471	9.51
192	2010	18.711	18.69	2010	9.493	9.7
193	2011	18.69	19.69	2011	9.543	9.52
194	2012	18.804	19.99	2012	9.554	9.51
195	2013	18.984	20.45	2013	9.548	9.61
196	Mean	17.94007027			8.397232432	
197	=CORREL(B12:B195,E12:E195)					

Due to their linear relationship, we can estimate local temperatures off the global temperatures. Here is a set of global temperatures which estimate the local temperature based off their linear relationship.

Mean Global °C	Mean Local °C	Ratio
6.000	12.816	2.136
6.250	13.350	2.136
6.500	13.884	2.136
6.750	14.418	2.136
7.000	14.952	2.136
7.250	15.486	2.136
7.500	16.020	2.136
7.750	16.554	2.136
8.000	17.088	2.136
8.250	17.622	2.136
8.500	18.156	2.136
8.750	18.690	2.136
9.000	19.224	2.136
9.250	19.758	2.136
9.500	20.292	2.136
9.750	20.826	2.136
10.000	21.360	2.136

The Ratio column was calculated by mean global and local temperatures:

2.136 = 18.037 °C / 8.443 °C

FINAL CONCLUSION

Indeed, the Earth is getting warmer, based on the small data set we have. My local city's temperature correlates to global temperature changes.