Robert Grantham
Udacity Data Analyst Nanodegree, Project 1
5/4/2020

As the first project for Udacity's Data Analyst Nanodegree, I completed an analysis of provided weather data. This comparison generated insight that speaks not only volumes about my own community but also of the entire globe.

First, I queried the provided data with SQL. Ultimately, I concluded it best to coalesce the city and global data together according to year. I discovered the only available data for the city of Austin, TX to compare with global data spanned from 1820 to 2013. I queried the data accordingly:

```
SELECT g.year,
       g.avg_temp AS global_avg_temp,
       c.avg_temp AS city_avg_temp
FROM   global_data AS g
       LEFT JOIN city_data AS c
              ON g.year = c.year
WHERE  c.city = 'Austin';
```

Transferring the resulting data set into Excel, I began an exploratory analysis. Initially, I converted the temps to a 5-year moving average since the initial line chart was too volatile to interpret with any measure of comfort.[1] Afterwards, I moved to reconfiguring the temperatures.

I converted the temperature data from Celsius to Fahrenheit using the formula (=CONVERT "cell", "C", "F") and rounded the resulting decimal to the nearest 2 zeroes.[2] Exploring the results, I arrived at four observations:

1. Global temperature is increasing by an average of 1.36 degrees Fahrenheit every 100 years.
2. Austin stays on average, approximately 20 degrees Fahrenheit higher than the average global temperature over the course of the timeline.
3. From 2008 to 2013, Austin's average yearly temperature spiked by over 2 degrees Fahrenheit.
4. If historical trend continues without interruption or alteration, the global average temperature will reach 49.96 degrees Fahrenheit by 2030.[3]

Next, I viewed additional data, adding a column to include a favorite city, San Francisco:

---

[1] View the accompanying Excel workbook for calculations & other details.
[2] I left the temperatures of the finalized chart as Celsius, but I converted the temperatures to Fahrenheit for my own better grasp of the data and its insights.
[3] I performed a forecast out to the year 2030 primarily as an exercise. Predicting weather is of course nearly impossible let alone with the amount of data available.

```
SELECT    A.year,
          A.avg_temp gl_temp,
          B.avg_temp austin_temp,
          C.avg_temp sf_temp
FROM      global_data A
LEFT JOIN city_data B
ON        A.year = B.year
LEFT JOIN city_data C
ON        A.year = C.year
WHERE     B.city = 'Austin'
AND       C.city = 'San Francisco';
```

The resulting chart revealed a temperature spike in San Francisco as well. From 2008 to 2013, its average yearly temperature jumped by 2.12 degrees Fahrenheit. Curiosity led me to check 2008 – 2013 data with another favorite city, Berlin, Germany.

Berlin's average yearly temperature decreased by 0.97 degrees over the same time span. What other factors led to the tremendous spike seen in the US? What other insights are there to be found?

I thoroughly enjoyed working on this project and I look forward to the next project with much enthusiasm!

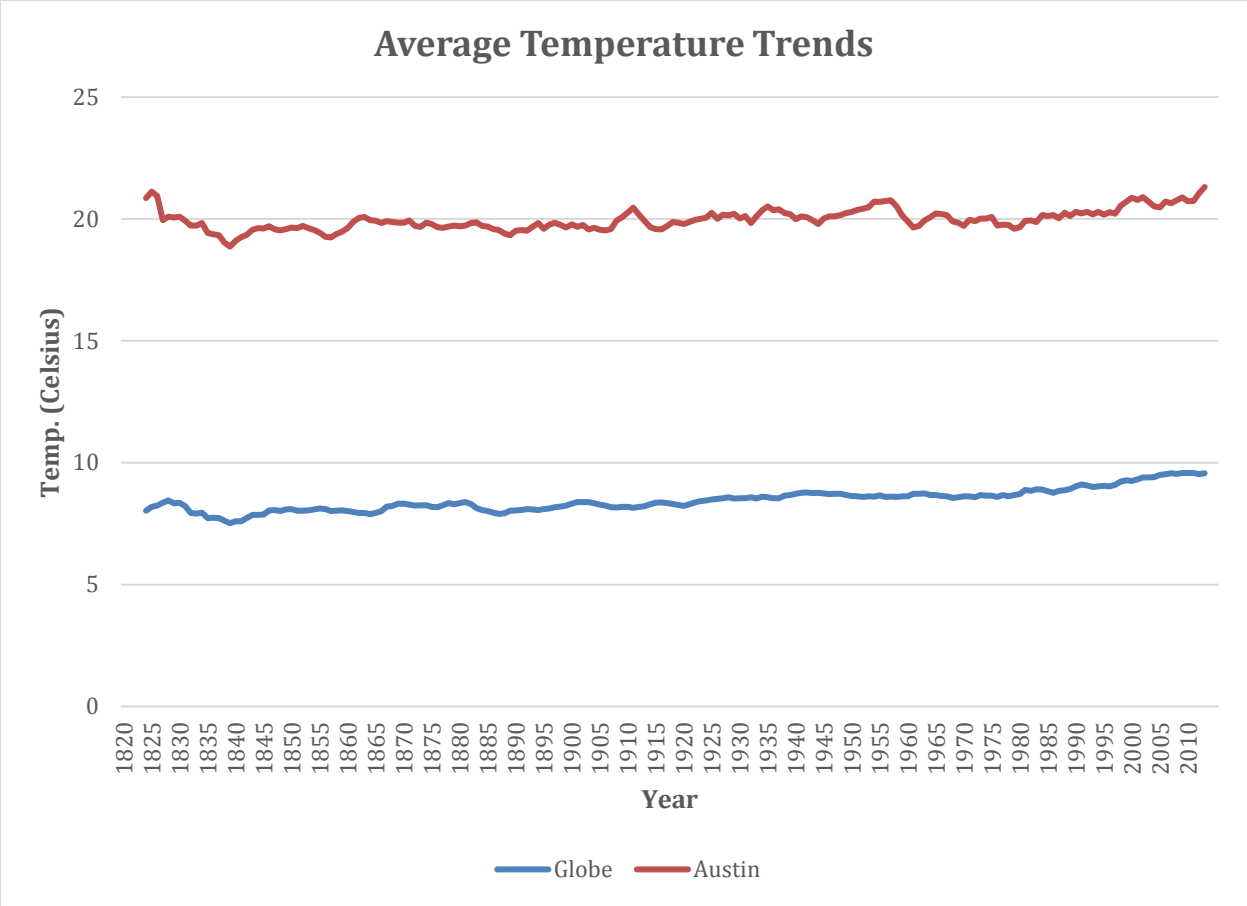
Average Temperature Trends
25
20
15
10
5
0
Temp. (Celsius)
1820
1825
1830
1835
1840
1845
1850
1855
1860
1865
1870
1875
1880
1885
1890
1895
1900
1905
1910
1915
1920
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
Year
Globe
Austin

Figure 1