Project 1 Submission – Kevin Wu

STEPS OUTLINE

1. I used SQL to determine if there was data for my city of interest, which was Santo Domingo, Dominican Republic. I used an SQL query to search by using a WHERE = clause.
2. I then merged the global_data table with a filtered city_data table by using this SQL query:

```sql
WITH sd1 AS (SELECT *
  FROM city_data
  WHERE city = 'Santo Domingo' AND country = 'Dominican Republic')

SELECT gd.year, gd.avg_temp as gl_temp, sd1.avg_temp as local_temp
FROM global_data as gd
JOIN sd1
ON sd1.year = gd.year;
```

3. I downloaded the .csv data.
4. I laid out the .csv data into Excel.
5. Due to the 5 consecutive years of missing data for the local temperatures between 1833 and 1837, I decided that a 10-year moving average would be best as to smooth over the missing data.
6. I also tested 2 and 3-year consecutive years, but I did not find it to be any more insightful than a 10-year moving average.
7. For data visualization, as per the assignment, I created a line chart. For better readability of the temperature changes, I changed the Y-axis to span 5-30C and made the minor gridlines to be 0.5C each for granularity since the temperature changes are small.

OBSERVATIONS

1. The trends between local and global temperatures trend very similarly.
2. Sometimes the magnitude of temperature change will be more or less when compared to the global temperature changes. However, in the end, the local temperatures seem to still reflect the global trend.
3. It does not appear that the big temperature difference between global average temperatures (around 8C) and local average temperatures (around 26C) affect the temperature change trend.
4. For both data, annual temperatures have been increasing. Local temperatures have increased 1C and global temperatures have increased 1.5C since 1833.

Project 1 Submission – Kevin Wu