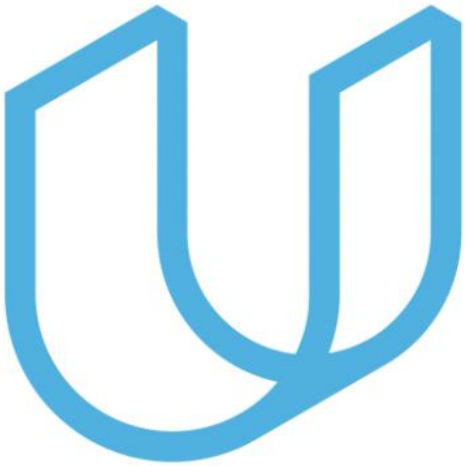
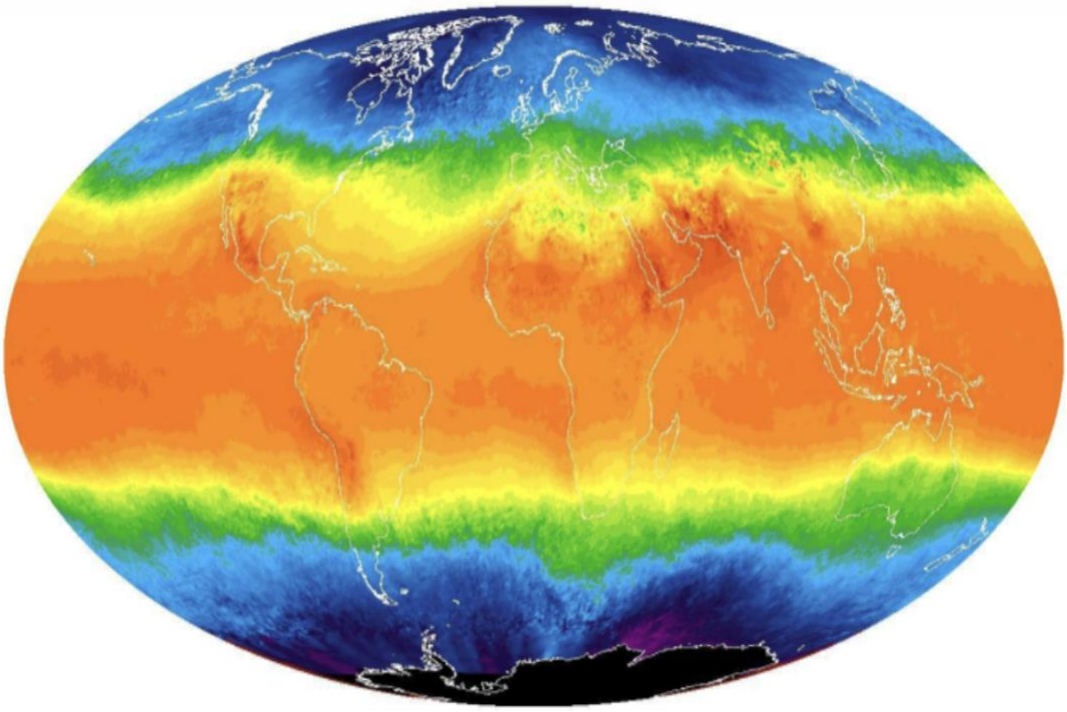
DATA ANALYST NANO DEGREE



UDACITY



Explore Weather Trends

By
Ramya N

Contents

Project Summary

In this project, I analyzed the temperatures of Detroit (MI), temperatures of Chicago (nearby city) and global temperature data and then compared the temperature trends.

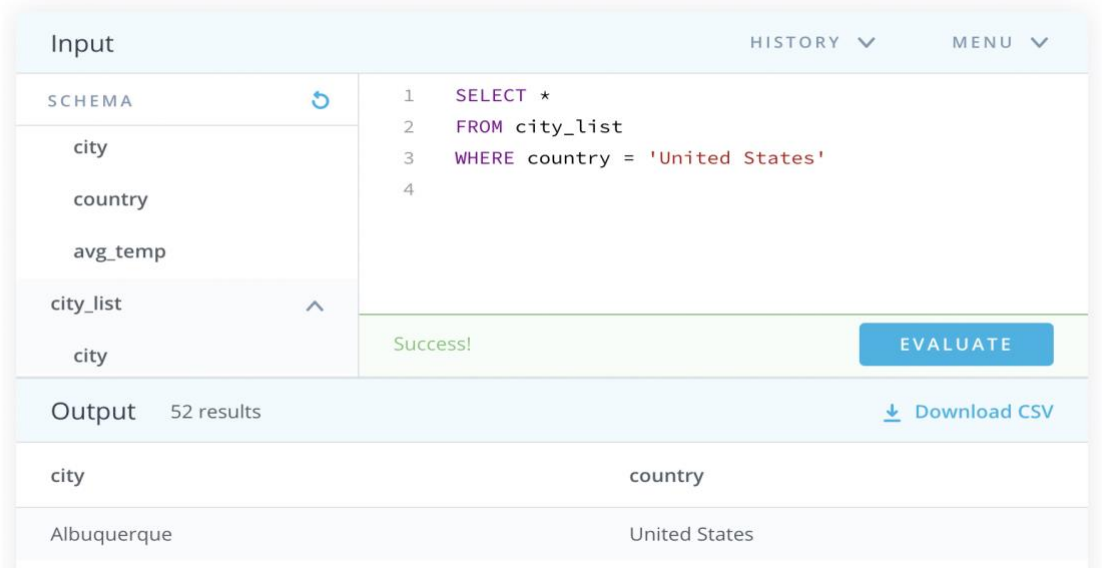
Introduction

- SQL Query was used to extract the data from given database.
- Once I get the data, then I downloaded the csv file using Excel.
- Created a line chart that compares Detroit (MI) temperatures with the global temperatures as well as with the Chicago city and also calculated moving averages (MA), Correlation Coefficient, Temperature estimation in Detroit.
- I generated five conclusions based on the analysis.

Analysis

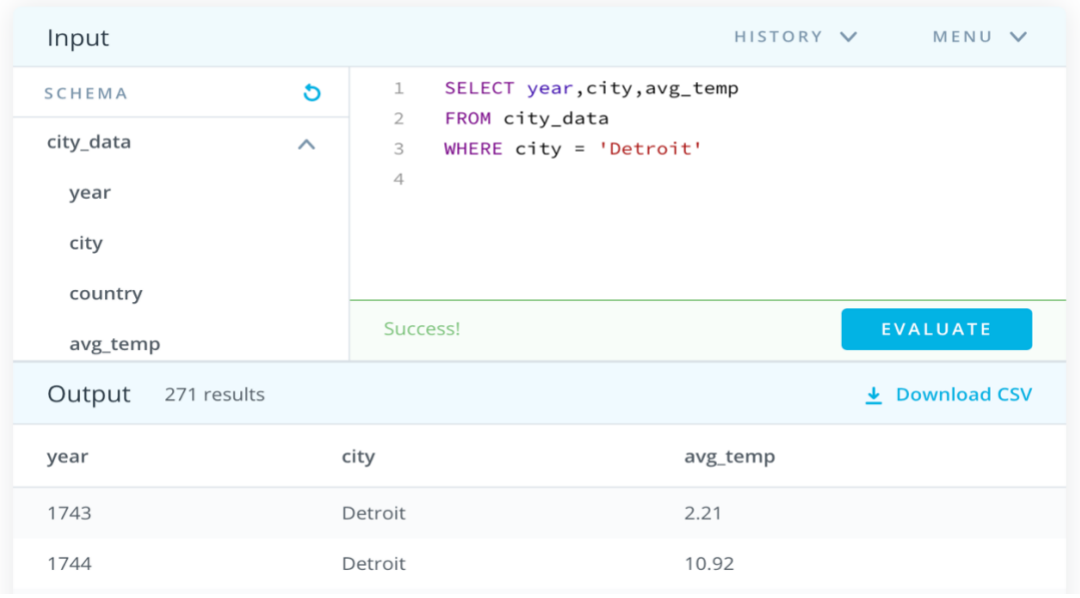
Step -1:

First step is to find the city which is closer to where I live. For that I have written a SQL Query to retrieve the cities in the united states.



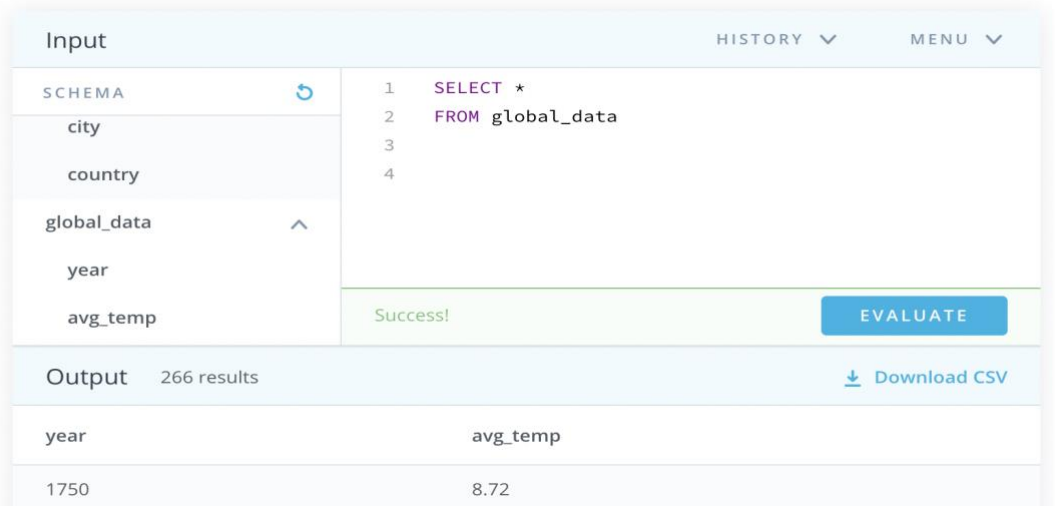
Step -2:

To extract the city level data, the following SQL Query was used.



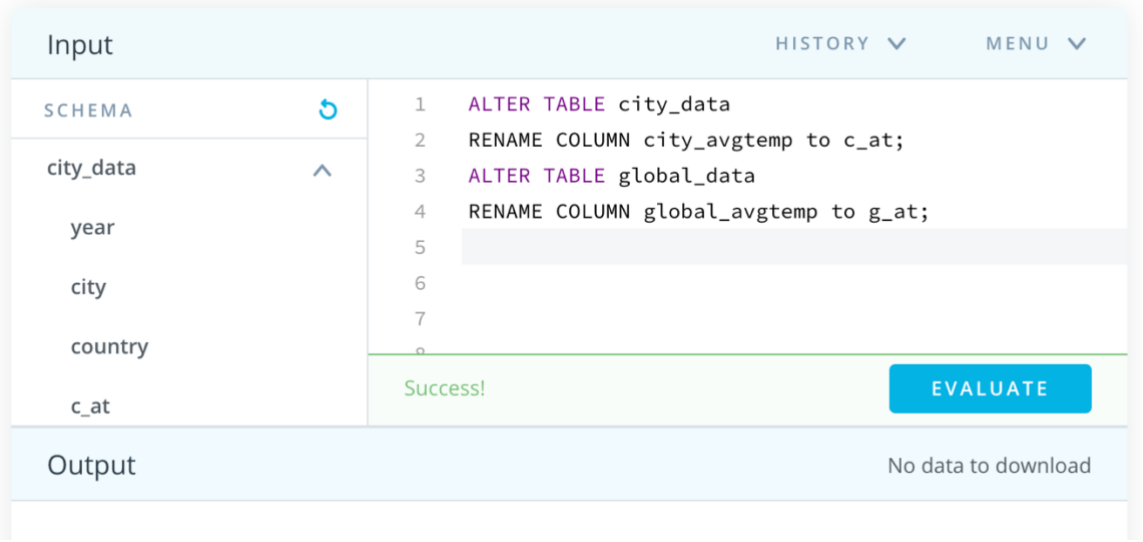
Step – 3:

To extract the global temperature data, the following SQL Query was used.



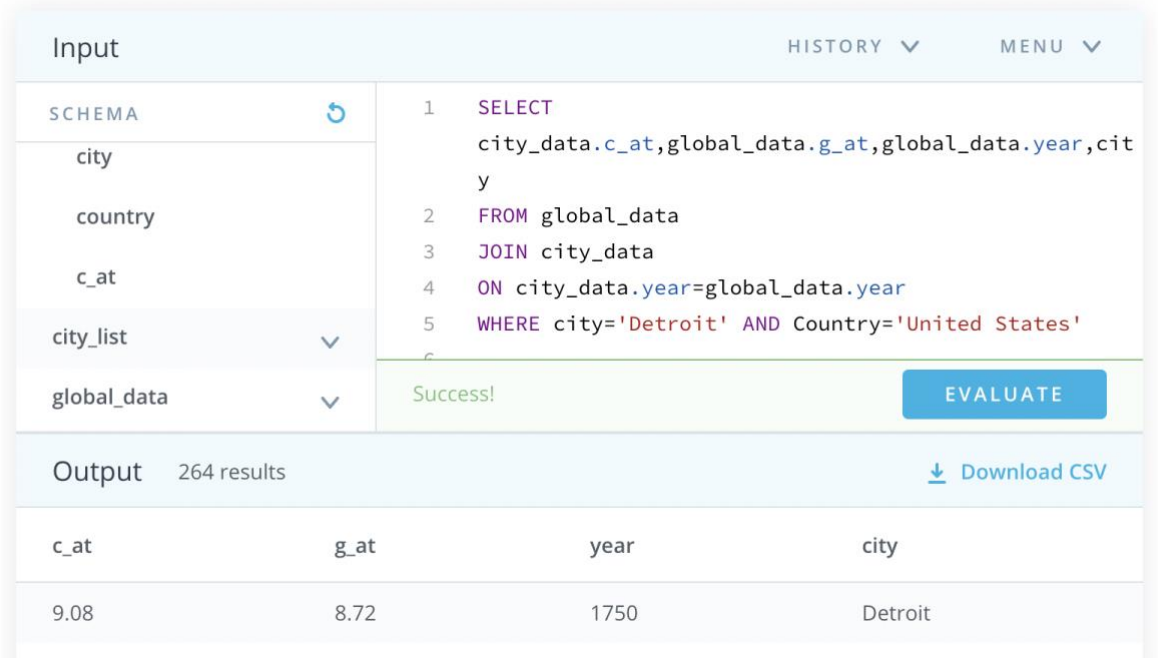
Step -4:

Renaming avg_temp column headers on two databases using SQL.



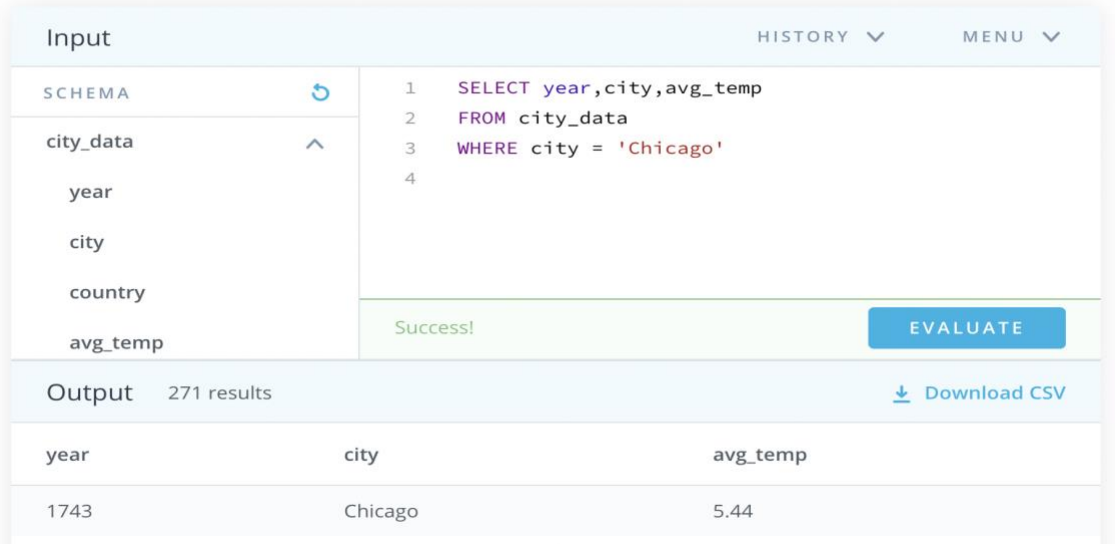
Step – 5:

I have joined the two tables using JOINS as the year is same in both tables.



Step – 6:

Extracted one more city (Chicago) to compare with Detroit and Global temperature trends.



Step – 7:

Downloaded the CSV file and used Microsoft Excel for the analysis.

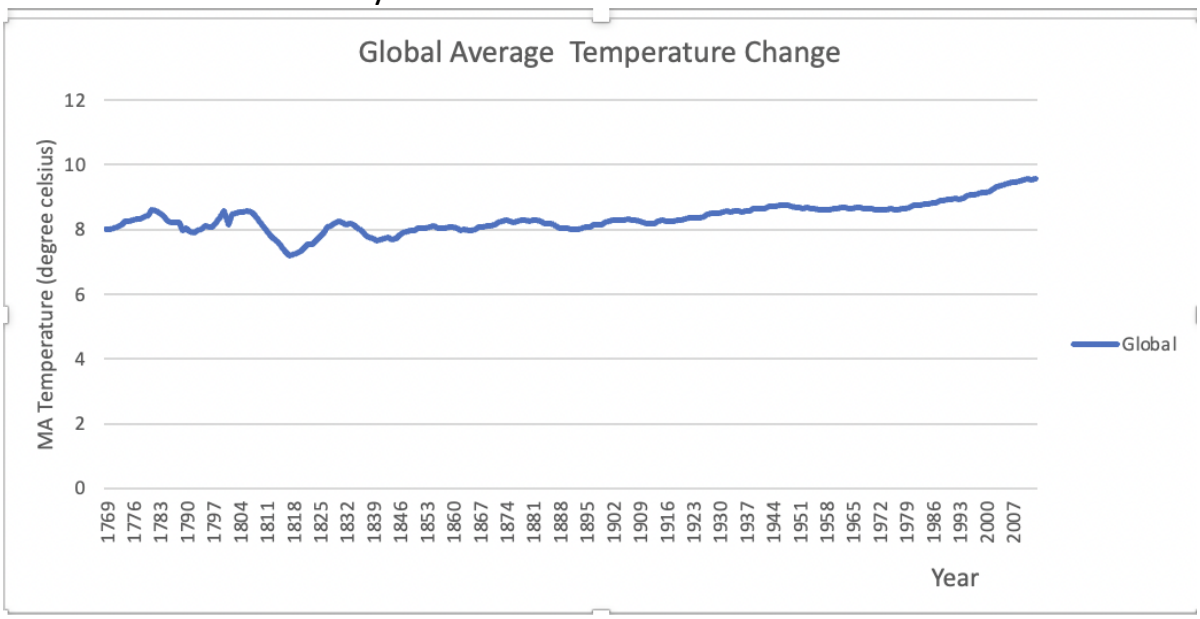
Moving Averages using Excel:

- Calculated Moving Average to smooth out data and also to observe the trends.
- In order to get the smooth line chart, I have used 10 year Moving Average.
- To get the moving average value for 10 years, I used a formula AVERAGE (A2:A11).

Plots and Data Visualization:

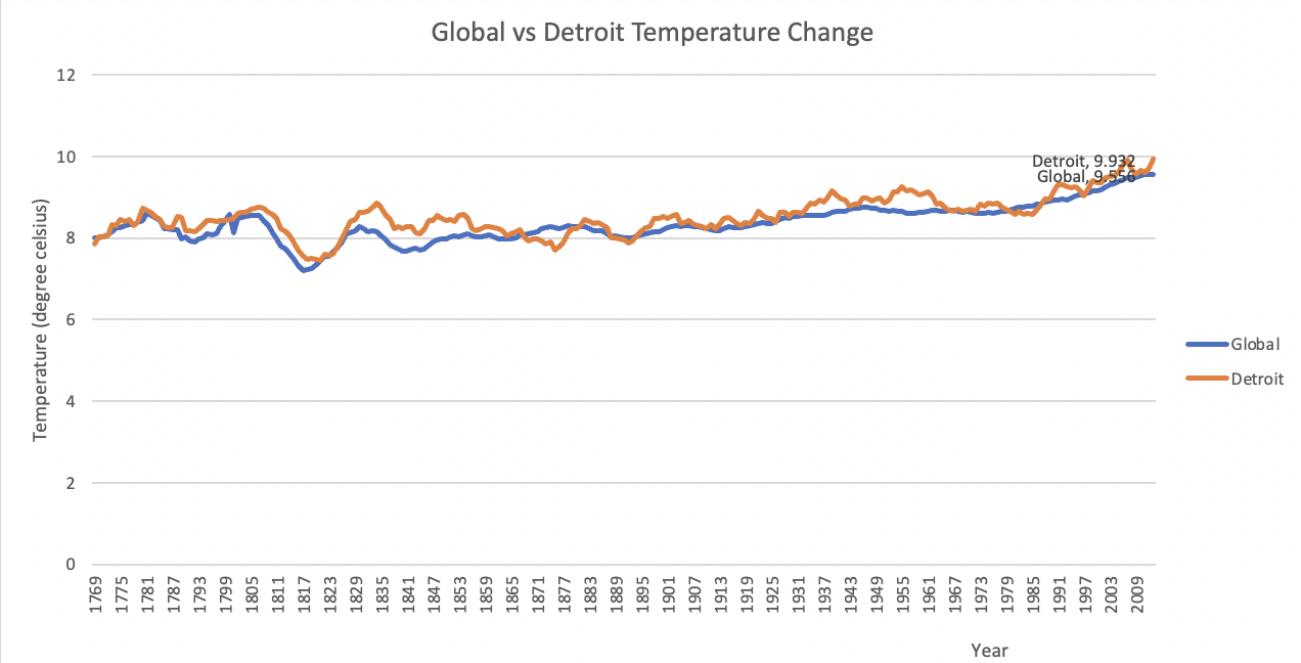
Step – 1:

Drawn a Line Chart for Global Average Temperature just to see if there is any difference between Detroit City and Global Data.



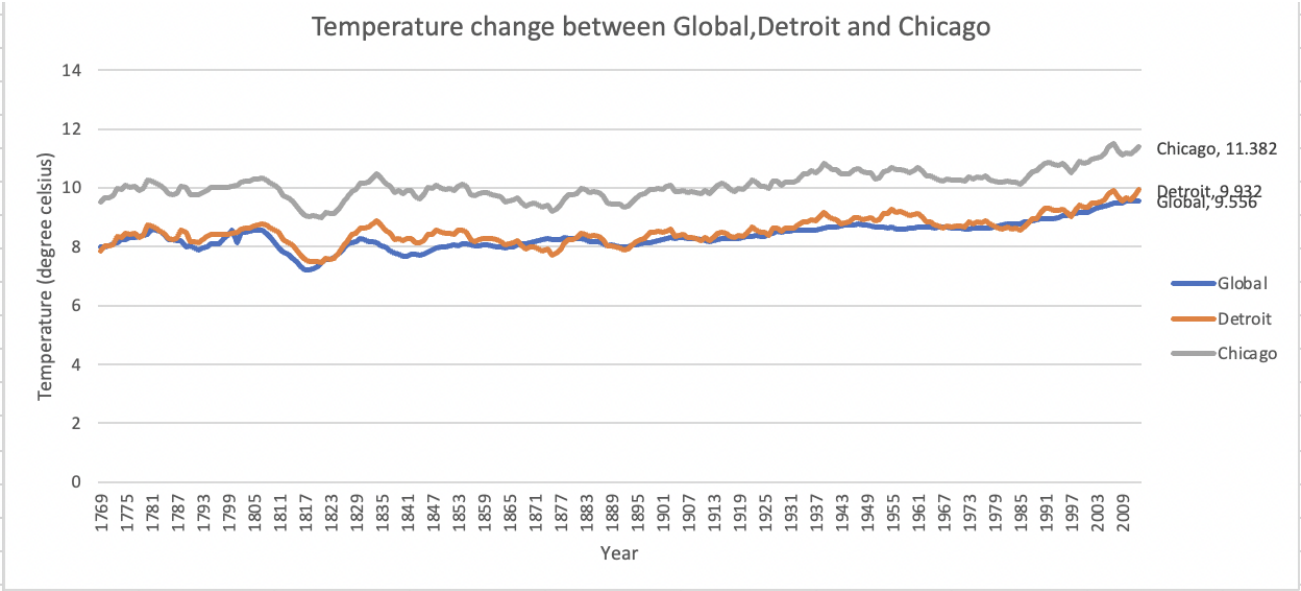
Step – 2:

Now Plotted line chart for both Global and Detroit Average Temperatures with 10 year Moving Average.



Step – 3:

Now added one more city (Chicago) to the above chart to observe the differences between Global, Detroit and Chicago average temperatures.



Observations:

- Detroit city observed to have temperature slightly greater than the global temperature.
- Whereas Chicago city has temperature greater than the global average.
- There was a slight decrease in temperature in 18th century, but it was there for a short time and then it increased in early 19th century and continued over the time with less fluctuations.
- According to the above charts, it shows that the temperature is raising over the years due to global warming.
- As well as the average temperature is consistent among Detroit, Chicago and global over the time.
- Detroit temperature is increasing overtime compared to the changes in the global average.
- The overall trend is increasing, which says that the world is getting hotter.
- Despite ups and downs from year to year, the global average temperature and Detroit, Chicago average temperatures are rising over the last 200 years.
- Detroit and Chicago average temperatures are rising constantly and directly proportionate to each other.

Correlation Coefficient:

Step -1 :

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.87508155
R Square	0.76576772
Adjusted R S	0.76483822
Standard Err	0.23753408
Observations	254

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	46.4839145	46.4839145	823.855119	2.1564E-81
Residual	252	14.2184544	0.05642244		
Total	253	60.7023689			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.40136324	0.28287341	1.41887932	0.15716982	-0.155734	0.95846047	-0.155734	0.95846047
8.03	0.97153915	0.03384815	28.7028765	2.1564E-81	0.90487785	1.03820044	0.90487785	1.03820044

RESIDUAL OUTPUT

Observation	*redicted 7.54*	*Residuals*
1	8.0541771	-0.7721771
2	8.13092869	-0.9269287

- The correlation coefficient for global and Detroit temperatures is 0.765 which says both are strongly related.
- X is the global temperature and Y is the Detroit City Temperature.

Step -2:

Correlation Coefficient between Chicago temperatures and Global temperatures:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.88723702
R Square	0.78718953
Adjusted R S	0.78631015
Standard Err	0.2206711
Observations	244

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	43.5905765	43.5905765	895.162092	2.8068E-83
Residual	242	11.784368	0.04869574		
Total	243	55.3749445			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2.14936512	0.26575663	8.0877196	2.9298E-14	1.62587369	2.67285656	1.62587369	2.67285656
7.995	0.95007604	0.03175466	29.9192596	2.8068E-83	0.88752522	1.01262686	0.88752522	1.01262686

RESIDUAL OUTPUT

Observation	*redicted 9.50*	*Residuals*
1	9.75757406	-0.0925741
2	9.7756255	-0.1066255

- The correlation coefficient for global and Detroit temperatures is 0.78 which means both are strongly correlated.
- X is the global temperature and Y is the Chicago City Temperature.

Step – 3:

- Correlation Coefficient between Detroit and Chicago cities temperature is 0.98
- All these Regression analysis shows that Detroit, Chicago and Global temperatures are strongly correlated to each other which means one increases, the other also increases.
- X is the Chicago city temperature and Y is the Detroit city Temperature.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.99254809
R Square	0.98515171
Adjusted R S	0.98509035
Standard Err	0.05828908
Observations	244

Temperature Estimation in Detroit based on the average global temperature:

- I used Y=a + bX to find out Detroit city temperature for a specific year, where X is global average temperature.

$$Y = 0.401 + 0.973\ (9.556)$$
$$= 9.69$$

- There is 0.4 difference in the (original data 9.932) to (9.69 estimated data) Which means the linear regression is not completely perfect. Further I have to do residual analysis to check if the model makes sense and if it's still acceptable.

References:

https://www.excel-easy.com/examples/regression.html
https://mathbits.com/MathBits/TISection/Statistics2/correlation.htm