Exploring Weather Trends

Project Outline

What tools did you use for each step? (Python, SQL, Excel, etc)

1. SQL is used to extract the data
   SELECT city_data.year, city_data.avg_temp, global_data.avg_temp
   FROM city_data, global_data
   WHERE city_data.year = global_data.year
   AND NOT city_data.avg_temp is NULL
   AND city_data.city = 'Singapore'

2. CSV file is downloaded from SQL and Microsoft Excel is applied to import the CSV for further analysis.

3. Average() function is applied to calculate 10,15,20,25 year moving averages for Singapore average temperature and global average temperature.

4. 25-year moving averages is chosen for both Singapore and global average temperature, as the line is smoothened without much noise.

5. Line graph is plotted using Microsoft Excel for further analysis.

6. Analysis Toolpak Add-ins is enabled in Microsoft Excel to conduct regression analysis.

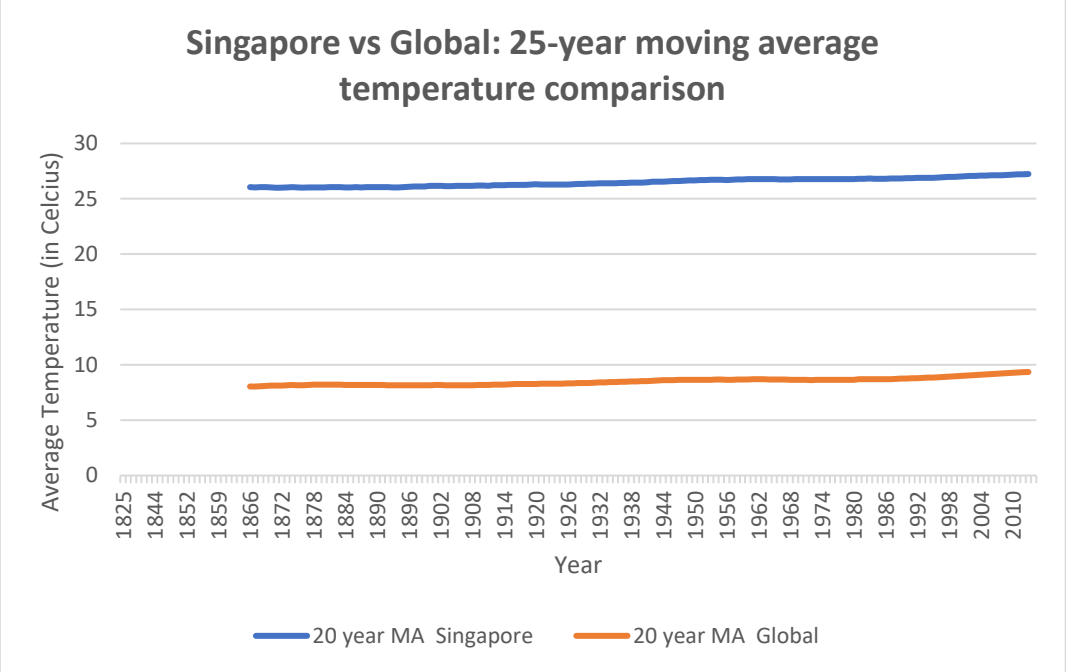
How did you calculate the moving average?

1. The moving average is calculated using Average() function in Microsoft Excel.

What were your key considerations when deciding how to visualize the trends?

1. Line graph
   a. 10,15,20 and 25 year moving average are calculated and the line graphs are plotted. It is observed that the 25 year moving average is the most suitable one for this data as the line is smoothened, without much noise.

Results and Interpretation



1.1 Line chart with Singapore and global temperature trends

1. The average temperature in Singapore is higher than average global temperature.

2. From the graph, we can see the difference in temperature has been consistent (for a difference at around 20 degree Celsius) since year 1825 to year 2013.

3. The graph above shows that the Singapore average temperature increases as the global temperature increases. This is to believe that both have a positive relationship.

4. The graph shows an upward trend in both Singapore and global average temperature. This indicates that the world is getting warmer.

5. The consistent upward trend is observed in both Singapore and global average temperature. From the graph, the average temperature in Singapore gradually increased to around 28 degree Celsius from 25 degree Celsius; and global average temperature rose from 8 degree Celsius to around 10 degree Celsius.

Regression Analysis

Regression Statistics	
Multiple R	0.975583965
R Square	0.951764073
Adjusted R Square	0.95143369
Standard Error	0.071855905
Observations	148

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	14.87429748	14.87429748	2880.78953	5.19593E-98
Residual	146	0.753837589	0.005163271		
Total	147	15.62813507			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-14.55484409	0.429614908	-33.87881525	4.7505E-71
X Variable 1	0.870131885	0.016211728	53.67298696	5.19593E-98

Linear Regression Equation:

*Singapore Average Temperature = 0.8701 * Global Average Temperature - 14.5548*

Intercept and X-variable 1 shows p-value of 0.000 (less than 0.05), which indicates the independent variable is significant.

Multiple R:

The Correlation Coefficient is 0.9755, which shows strong positive relationship.
This implies that Singapore average temperature increases when global temperature increases

R-Square:

The r-square is 0.9518, which indicates that 95% of the dependent variables (global temperature) are explained by the independent variable.
Thus, this model is considered a good fit.

Standard Error:

The standard error is 0.07186 or 7.186%, which is relatively low. This shows that the regression analysis has a high precision (above 90%)

From the ANOVA analysis, Significance F shows a value of 0.000 (less than 0.05), which indicates the model is a good model.

To conclude, the linear regression model is a good model to make prediction of the Singapore average temperature given global average temperature.