

Analysis and Comparing Global and Bangalore City Average Temperature Trend:

STEP-1:

I used the data base provided by the udacity which consists of three relations named **city_list**, **city_data**, **global_data** and extracted required data which is to be used for comparing trends.

after observing that number of rows in Global_data are more than number of rows in city_data for Bangalore. I extracted only data of years which are common in both city_data and global_data for accurate comparison rather than replacing missing data with mean value which is considered as best alternative for missing values.

QUERY USED:

SELECT city_data.year,city_data.avg_temp AS avg_city,global_data.avg_temp AS avg_global FROM city_data INNER JOIN global_data ON city_data.year=global_data.year WHERE city='Bangalore'

The screenshot shows the Udacity Data Analyst interface. The left sidebar contains a navigation menu with sections: SEARCH, RESOURCES, and CONCEPTS. Under CONCEPTS, there is a list of projects: 1. Your First Project, 2. Project Instructions, 3. Accessing Data With SQL (selected), 4. Moving Averages, and 5. Project: Explore Weather Trends. The main area is titled 'Accessing Data With SQL' and shows a query editor with the following SQL query:

```
1 select city_data.year,city_data.avg_temp as avg_city,global_data.avg_temp as avg_global from city_data inner join global_data on city_data.year=global_data.year where city='Bangalore'
```

The query is successful, and the output shows 218 results. The output table has three columns: year, avg_city, and avg_global. The first two rows of data are:

year	avg_city	avg_global
1796	24.49	8.27
1797	25.18	8.51

- downloaded the output of my query as a csv file.
- Imported the csv file to a google sheet.

Untitled spreadsheet

year	avg_city	avg_global
1796		8.27
1797		8.51
1798		8.67
1799		8.51
1800		8.48
1801		8.59
1802		8.58
1803		8.5
1804		8.84
1805		8.56
1806		8.43
1807		8.28
1808		7.63
1809		7.08
1810		6.92
1811		6.86
1812		7.05
1813	24.23	7.74
1814	23.91	7.59
1815	23.79	7.24

STEP-2:

I Replaced the missing values in present dataset by the mean value of its column.

function used to replace missing values by mean values is :

=IF(ISBLANK(A2),ROUNDUP(AVERAGE(A2:\$A219),2),A2) and applied it to whole average temperature columns

Weather Trends

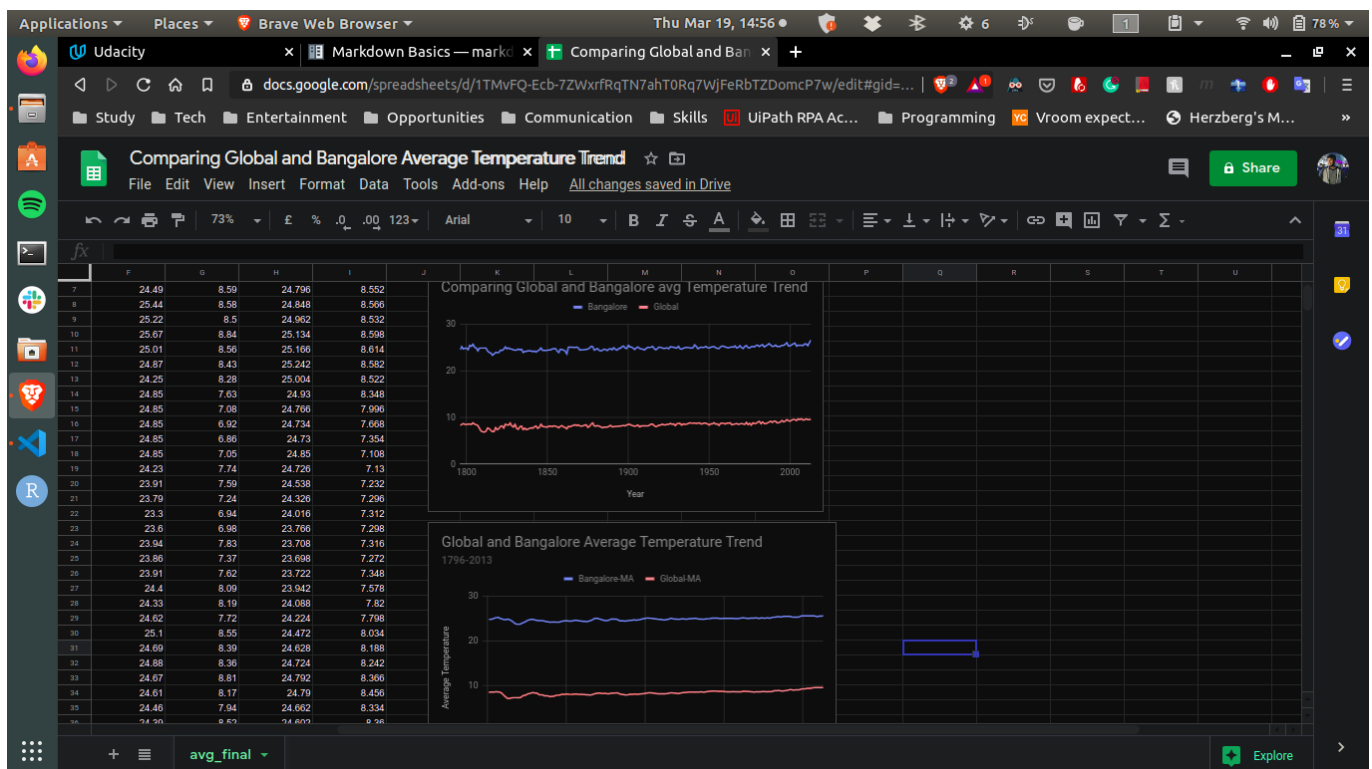
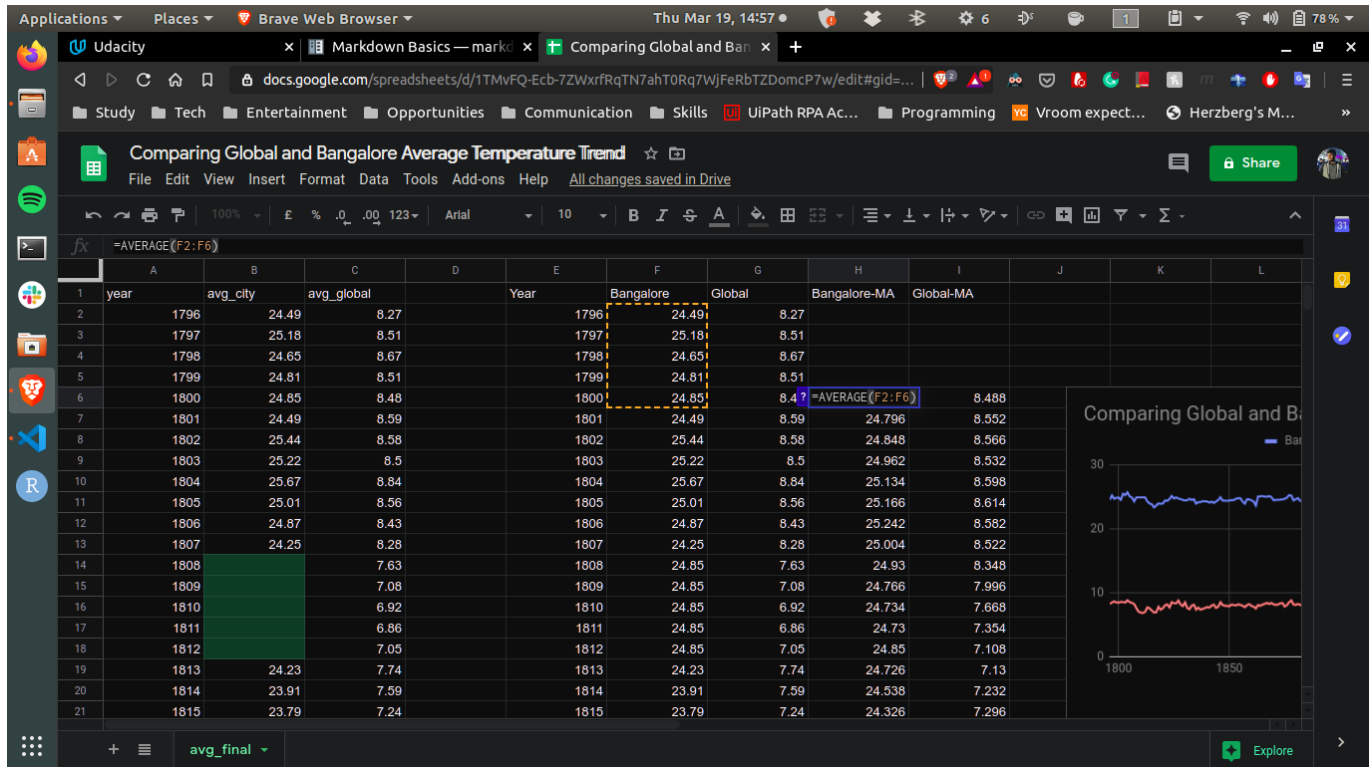
year	avg_city	avg_global
1796		8.27
1797		8.51
1798		8.67
1799		8.51
1800		8.48
1801		8.59
1802		8.58
1803		8.5
1804		8.84
1805		8.56
1806		8.43
1807		8.28
1808		7.63
1809		7.08
1810		6.92
1811		6.86
1812		7.05
1813	24.23	7.74
1814	23.91	7.59
1815	23.79	7.24

STEP-3:

I created a line chart comparing Global and bangalore city average temperatures. In order to smooth out the lines and make trends more observable I computed 5-year Moving Averages.

FUNCTION USED:

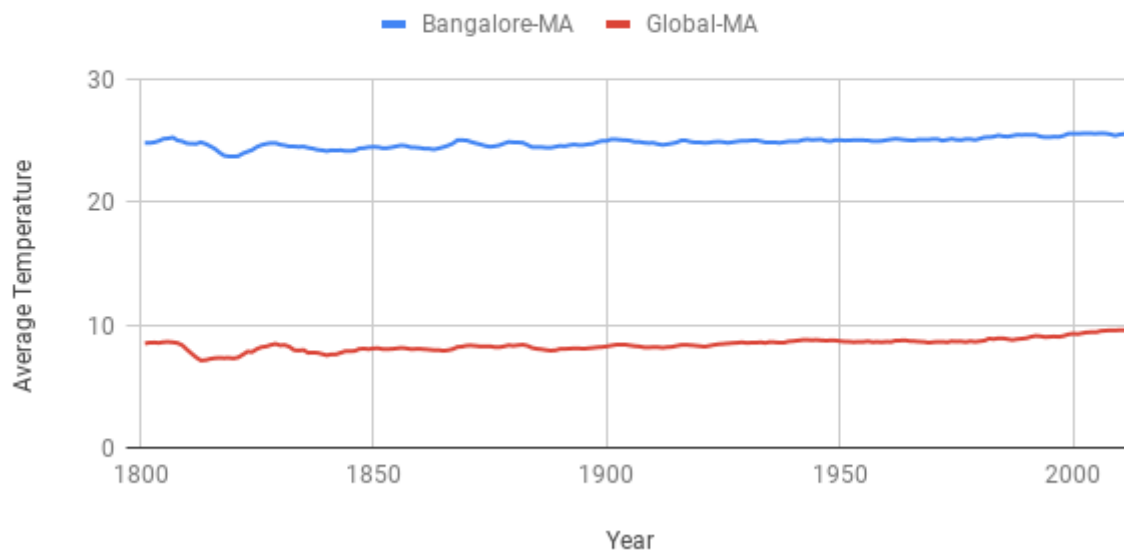
=AVERAGE(F2:F6) and applied it to both of the average temperature columns.



FINAL CHART:

Global and Bangalore Average Temperature Trend

1796-2013



OBSERVATIONS:

- The biggest difference between "Bangalore-MA" and "Global-MA" is **17.7** at **1812**.
- The Global and Bangalore average temperature trends have a positive correlation coefficient of **0.79**.
- Least Global and Bangalore average temperatures are **6.86** and **23.3** respectively.
- Mean Global and Bangalore average temperatures are **8.40** and **24.85** respectively.
- For every increase of **100** in Year, "Global Moving Average" increases by about **0.674**.