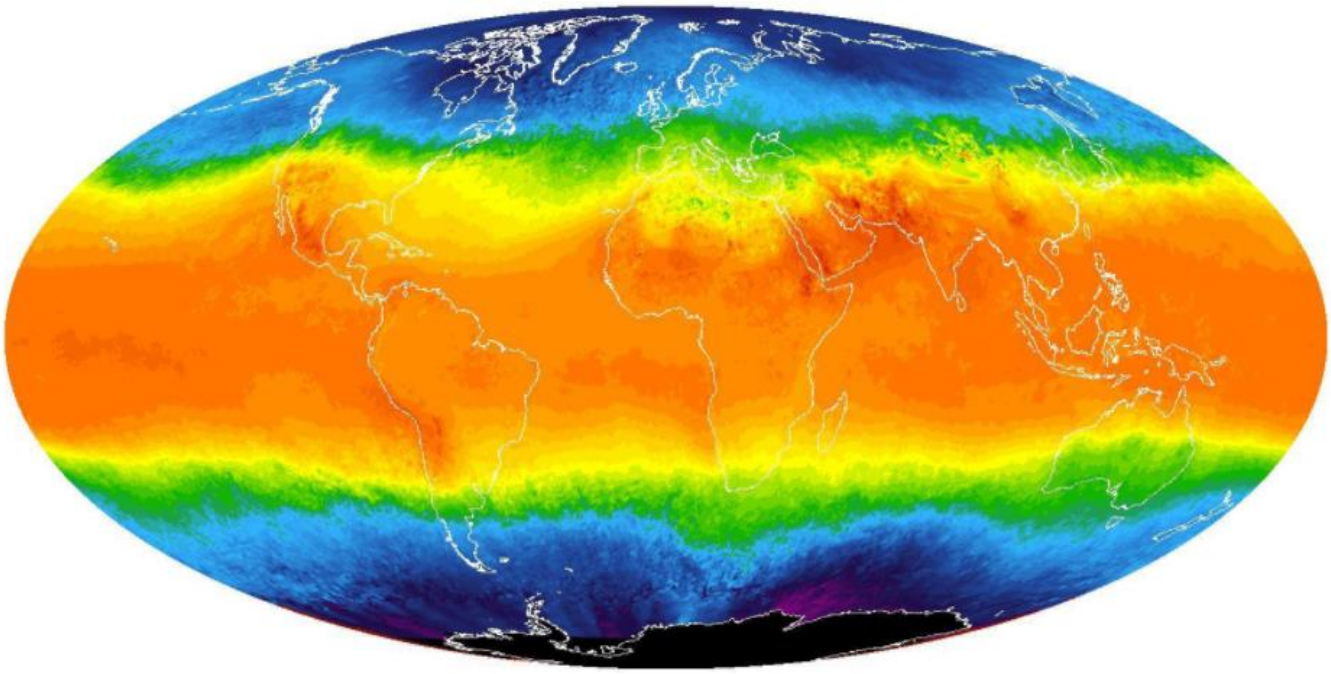
Project: Explore Weather Trends



Report Submitted By:
Avinash Korwarkar

Project: Explore Weather Trends

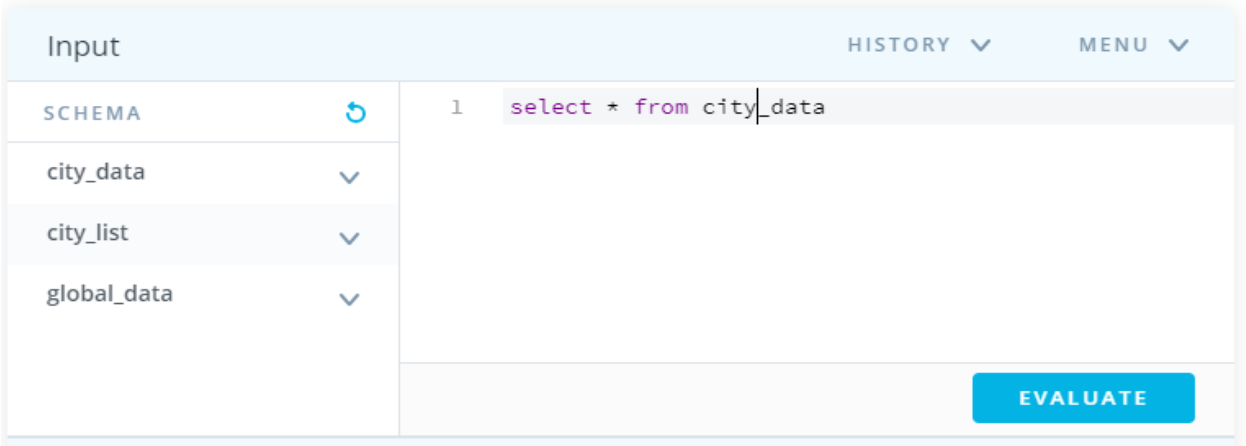
The project involves extracting data on world and city average temperatures over several years, analyze by visualization and make inferences on overall weather trends.

Below are the steps followed in completing the project.

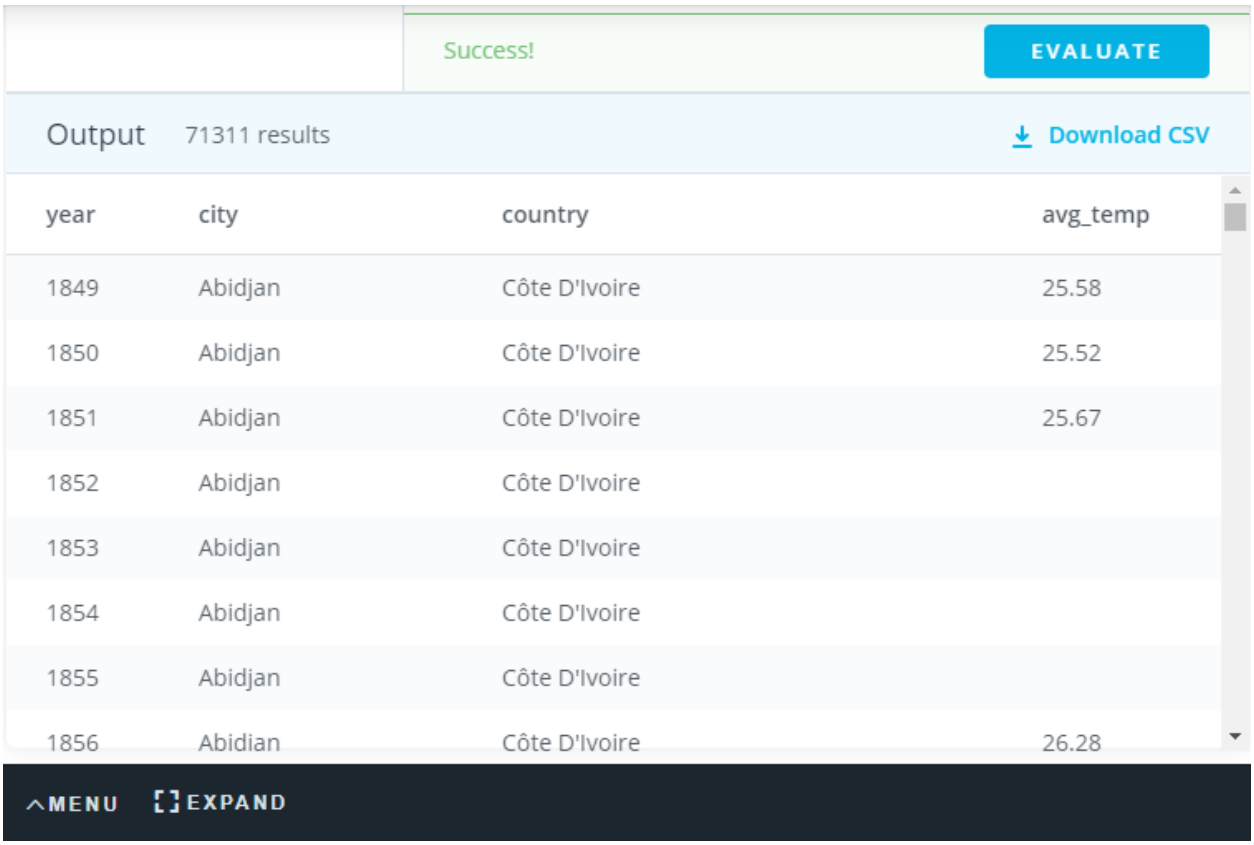
1) Extracting data:

Tool Used: *SQL*

We need to run the below SQL query to extract all the data in the city_data table.

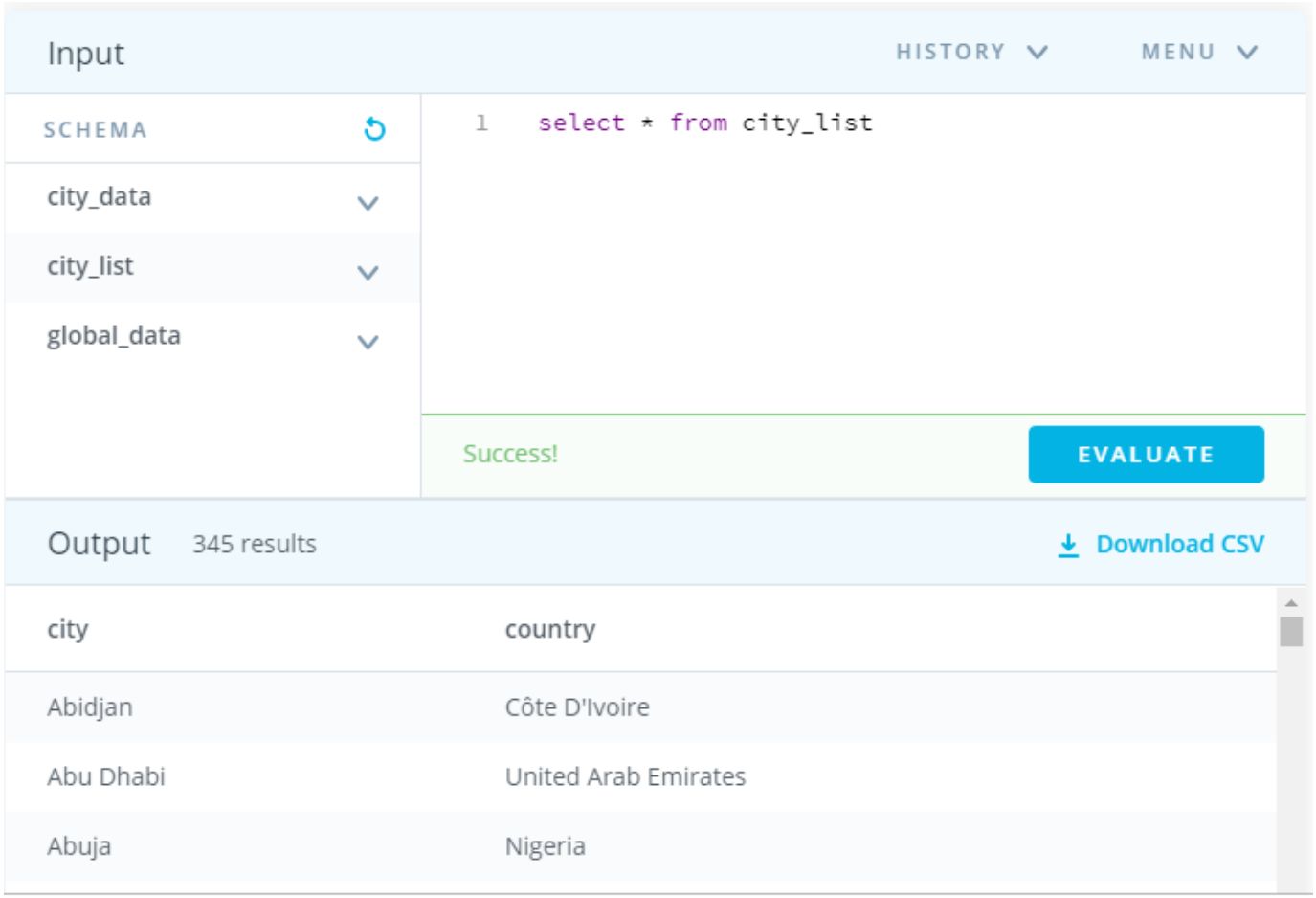


We then click on the EVALUATE button to run the query. If the query is correct, the data requested in the query is extracted and a “Success” message is displayed as below.



The extracted data can then be downloaded to a .csv file by clicking on the Download CSV link. Once the data downloads, a link to the extracted data is available at the bottom of the window.

Similarly, we can download data from the city_list and global_data tables as below.



Input HISTORY MENU

SCHEMA

city_data

city_list

global_data

```
1  select * from global_data
```

Success! EVALUATE

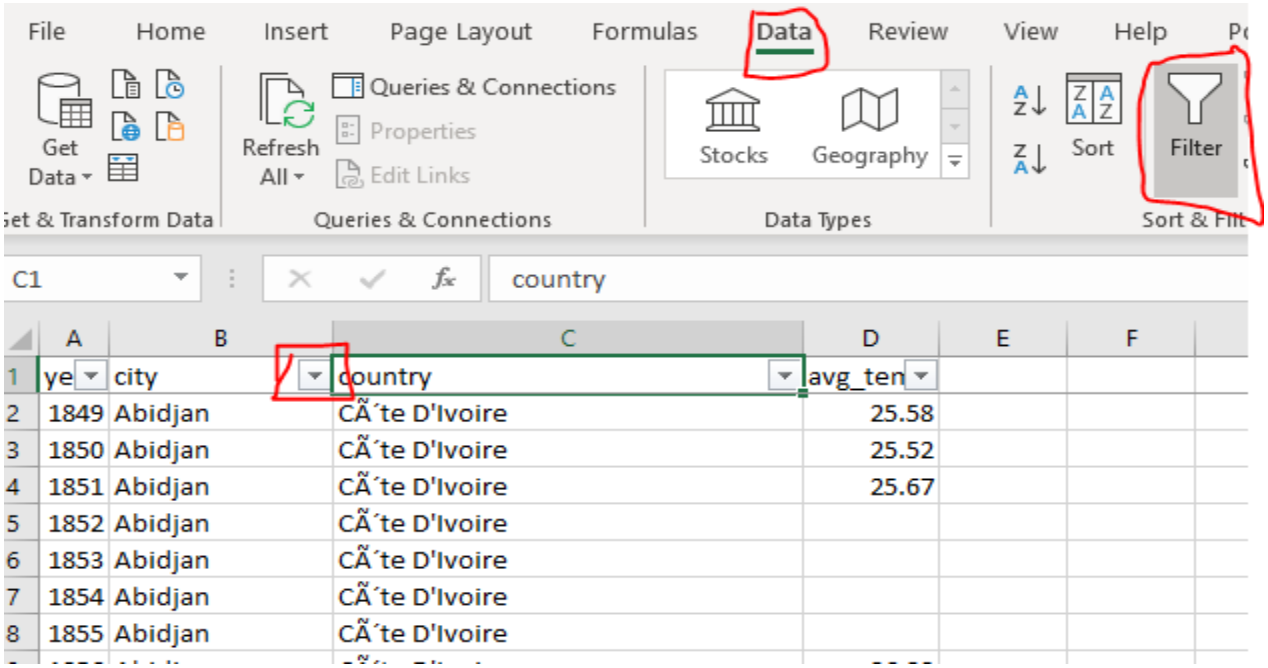
Output 266 results Download CSV

year	avg_temp
1750	8.72
1751	7.98
1752	5.78

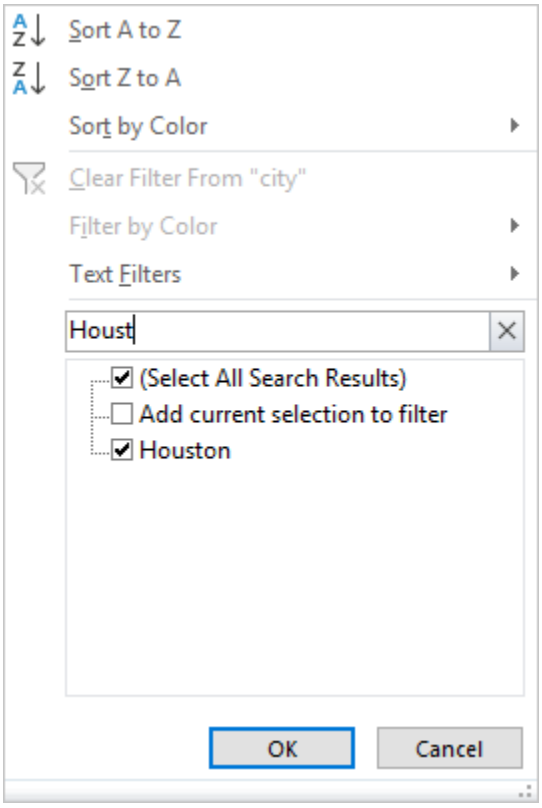
2) **Preparing data for analysis / visualization:**

Tool Used: *Excel*

The .csv file for city_data contains average temperatures for several cities in the world. I am from Houston, TX and need to extract data for Houston from the larger set. We can do this by filtering out data for Houston using using Excel's "Filter" function.



Clicking on the drop down provides the option to filter for Houston in the “city” column as below.



After checking Houston and clicking OK, the below data gets extracted from the larger set. This data can then be copied on to another worksheet and named “Houston”.

	A	B	C	D
1	ye	city	country	avg_ten
25954	1820	Houston	United States	19.11
25955	1821	Houston	United States	19.57
25956	1822	Houston	United States	20.05
25957	1823	Houston	United States	19.62
25958	1824	Houston	United States	20.19
25959	1825	Houston	United States	20.44
25960	1826	Houston	United States	20.17
25961	1827	Houston	United States	20.83
25962	1828	Houston	United States	20.41
25963	1829	Houston	United States	20
25964	1830	Houston	United States	20.72
25965	1831	Houston	United States	19.25
25966	1832	Houston	United States	19.89
25967	1833	Houston	United States	20.32
25968	1834	Houston	United States	20.56
25969	1835	Houston	United States	18.62
25970	1836	Houston	United States	19.01

Our goal is to compare Houston's average temperature with average global temperature. So we will need to extract the average global temperature for a year from the "global_list" sheet against the average Houston temperature for that year. This can be done using Excel's "vlookup" function.

"**city_data**" sheet

E2 =VLOOKUP(C2,global_data!A:B,2,0)

	A	B	C	D	E	F	G
1	city	country	Ye	Housto	World	Houston-10YMA	World-10YMA
2	Houston	United States	1820	19.11	7.62		
3	Houston	United States	1821	19.57	8.09		
4	Houston	United States	1822	20.05	8.19		

"**global_data**" sheet

	A	B
1	year	avg_ter
2	1750	8.72
3	1751	7.98
4	1752	5.78
5	1753	8.39
6	1754	8.47

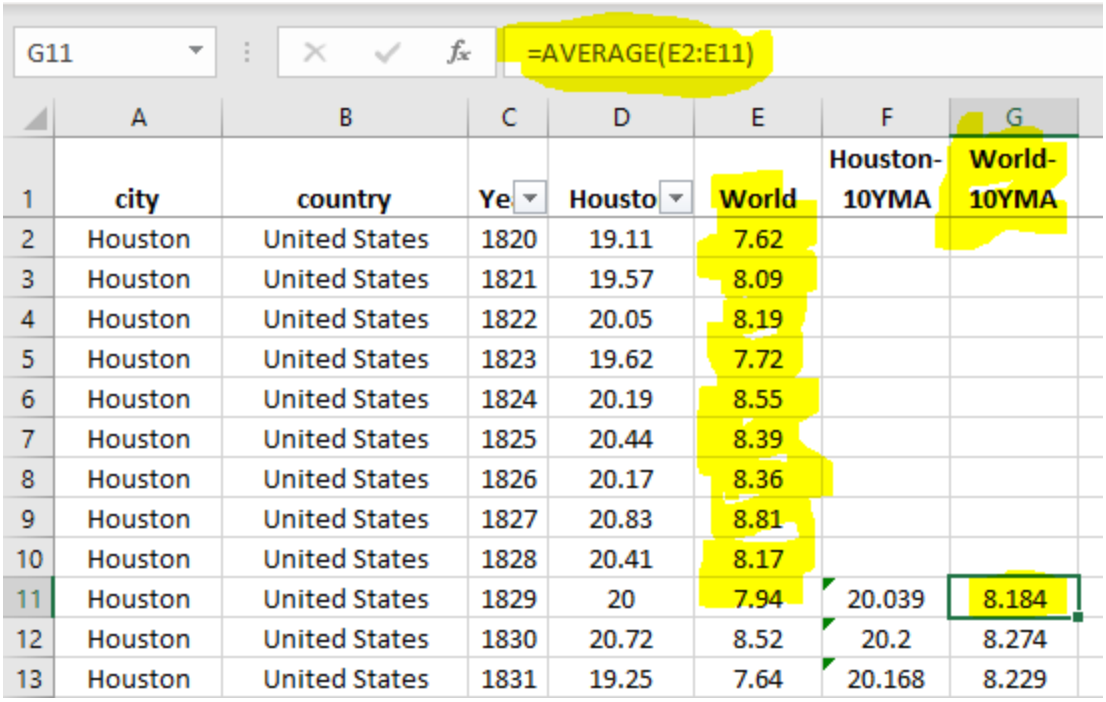
We can now plot a line chart for average global temperatures and average Houston temperatures on the Y-axis for various years on the X-axis, but the line is going to be **choppy** because of the year-to-year variation in average temperatures. We can smoothen this out by plotting the **moving average temperatures** instead.

3) Calculating Moving Averages:

I chose to calculate 10 year moving averages for both Houston and Global temperatures. This is done using Excel's "average" function as below.

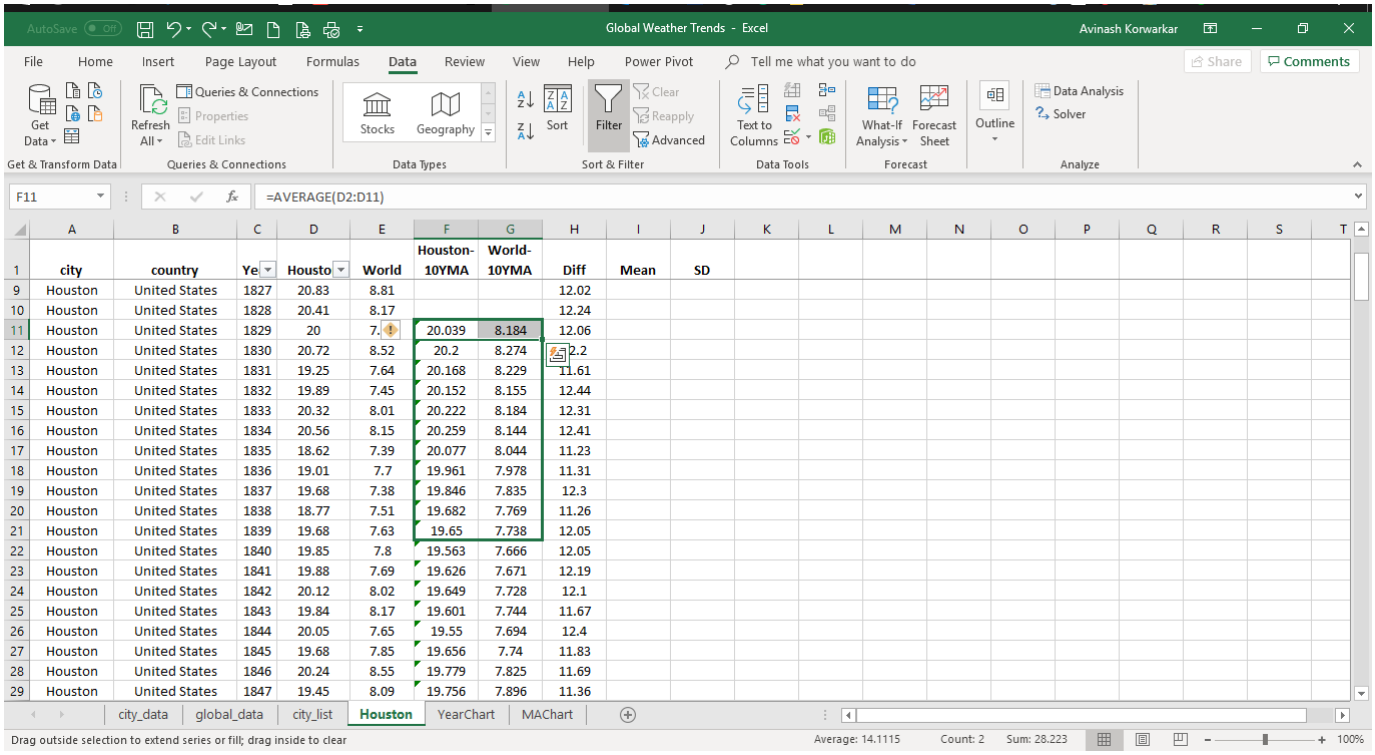
F11 =AVERAGE(D2:D11)

	A	B	C	D	E	F	G
1	city	country	Ye	Housto	World	Houston-10YMA	World-10YMA
2	Houston	United States	1820	19.11	7.62		
3	Houston	United States	1821	19.57	8.09		
4	Houston	United States	1822	20.05	8.19		
5	Houston	United States	1823	19.62	7.72		
6	Houston	United States	1824	20.19	8.55		
7	Houston	United States	1825	20.44	8.39		
8	Houston	United States	1826	20.17	8.36		
9	Houston	United States	1827	20.83	8.81		
10	Houston	United States	1828	20.41	8.17		
11	Houston	United States	1829	20	7.	20.039	8.184
12	Houston	United States	1830	20.72	8.52	20.2	8.274
13	Houston	United States	1831	19.25	7.64	20.168	8.229
14	Houston	United States	1832	19.89	7.45	20.152	8.155



	A	B	C	D	E	F	G
1	city	country	Ye	Housto	World	Houston-10YMA	World-10YMA
2	Houston	United States	1820	19.11	7.62		
3	Houston	United States	1821	19.57	8.09		
4	Houston	United States	1822	20.05	8.19		
5	Houston	United States	1823	19.62	7.72		
6	Houston	United States	1824	20.19	8.55		
7	Houston	United States	1825	20.44	8.39		
8	Houston	United States	1826	20.17	8.36		
9	Houston	United States	1827	20.83	8.81		
10	Houston	United States	1828	20.41	8.17		
11	Houston	United States	1829	20	7.94	20.039	8.184
12	Houston	United States	1830	20.72	8.52	20.2	8.274
13	Houston	United States	1831	19.25	7.64	20.168	8.229

The formula can be copied for the remainder of the data set to calculate moving averages for the entire year range.



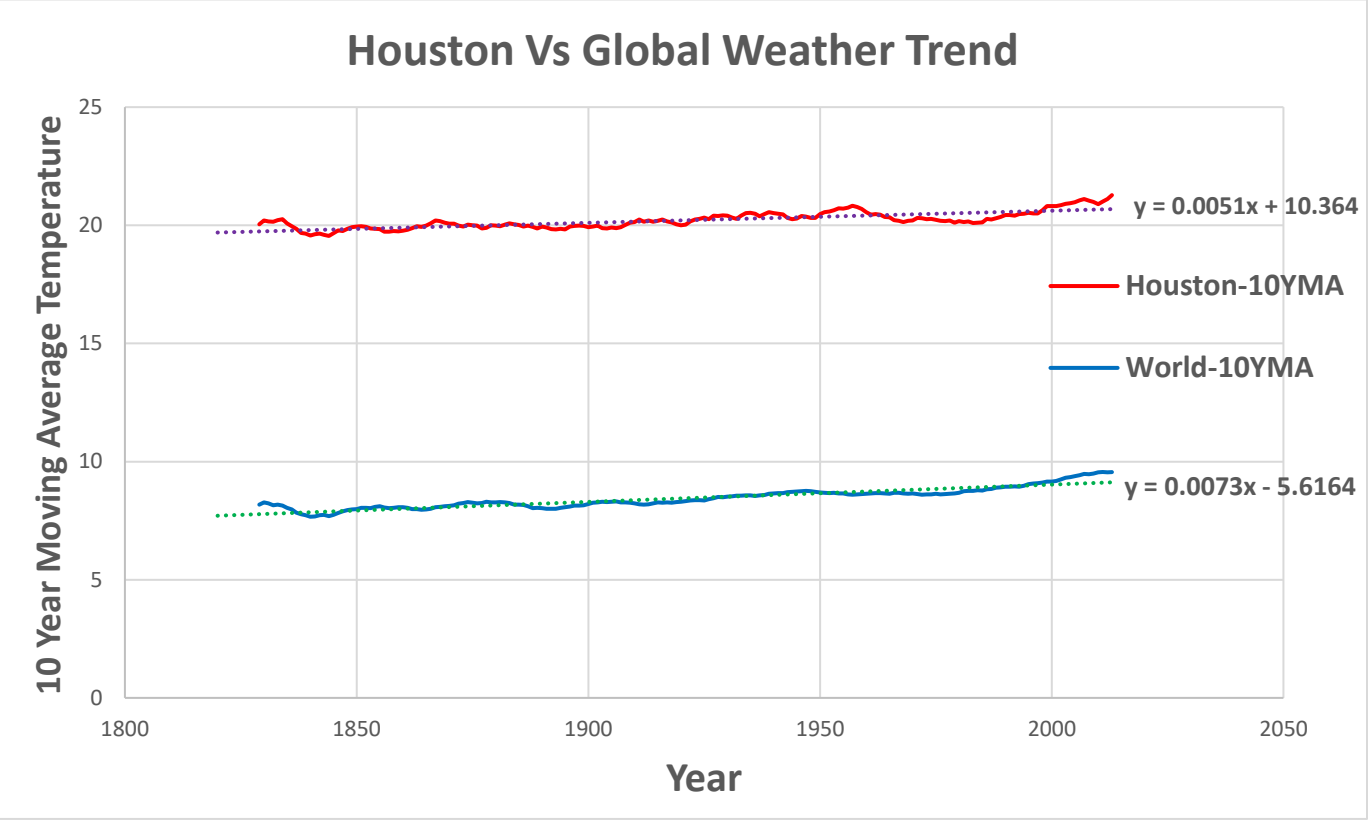
	A	B	C	D	E	F	G	H	I	J
1	city	country	Ye	Housto	World	Houston-10YMA	World-10YMA	Diff	Mean	SD
9	Houston	United States	1827	20.83	8.81			12.02		
10	Houston	United States	1828	20.41	8.17			12.24		
11	Houston	United States	1829	20	7.	20.039	8.184	12.06		
12	Houston	United States	1830	20.72	8.52	20.2	8.274	2.2		
13	Houston	United States	1831	19.25	7.64	20.168	8.229	11.61		
14	Houston	United States	1832	19.89	7.45	20.152	8.155	12.44		
15	Houston	United States	1833	20.32	8.01	20.222	8.184	12.31		
16	Houston	United States	1834	20.56	8.15	20.259	8.144	12.41		
17	Houston	United States	1835	18.62	7.39	20.077	8.044	11.23		
18	Houston	United States	1836	19.01	7.7	19.961	7.978	11.31		
19	Houston	United States	1837	19.68	7.38	19.846	7.835	12.3		
20	Houston	United States	1838	18.77	7.51	19.682	7.769	11.26		
21	Houston	United States	1839	19.68	7.63	19.65	7.738	12.05		
22	Houston	United States	1840	19.85	7.8	19.563	7.666	12.05		
23	Houston	United States	1841	19.88	7.69	19.626	7.671	12.19		
24	Houston	United States	1842	20.12	8.02	19.649	7.728	12.1		
25	Houston	United States	1843	19.84	8.17	19.601	7.744	11.67		
26	Houston	United States	1844	20.05	7.65	19.55	7.694	12.4		
27	Houston	United States	1845	19.68	7.85	19.656	7.74	11.83		
28	Houston	United States	1846	20.24	8.55	19.779	7.825	11.69		
29	Houston	United States	1847	19.45	8.09	19.756	7.896	11.36		

These moving averages are not as volatile as the individual temperatures and will provide for a smoother line chart.

4) Plotting the line chart

We need to now plot these moving averages **against the same year for both Houston and Global** temperatures. So the year needs to be on the X-axis and Houston's and Global moving average should be

on the Y-axis. This will also help us visually compare Houston's averages to Global averages since they are plotted next to each other on the same chart.



H2 =+D2-E2

	A	B	C	D	E	F	G	H	I	J
1	city	country	Ye	Housto	World	Houston-10YMA	World-10YMA	Diff	Mean	SD
2	Houston	United States	1820	19.11	7.62			11.49	11.765	0.508381
3	Houston	United States	1821	19.57	8.09			11.48		
4	Houston	United States	1822	20.05	8.19			11.86		
5	Houston	United States	1823	19.62	7.72			11.9		
6	Houston	United States	1824	20.19	8.55			11.64		
7	Houston	United States	1825	20.44	8.39			12.05		

5) **Observations and Inferences:**

- **a)** It can be clearly inferred from the line chart above that **Houston is definitely hotter** than the global average, with Houston's average hovering around **20 deg** and Global average hovering around **8.5 deg**.
- **b)** I added another column (H) for the difference in average temperature between Houston and Global for all the years and then calculated the mean and standard deviation of this difference. We can infer from the data that **Houston is 11.76 ± 0.51 degrees hotter** than average global temperatures.
- **c)** Since the standard deviation is fairly small (0.51) as compared to the average (11.76), we can assume that the **difference between Houston and Global averages has been fairly consistent** over the years, with very little variation.
- **d)** It can also be inferred that over a period of time that both Houston's average and Global average **temperatures are rising gradually**, since **the slope of the trend lines is positive.**
- **e)** The data also shows that this gradual increase follows a **linear trend and has been fairly consistent** over the last 200 years.

f) The slope of the trend line for Houston is 0.0051. This means that average temperature for Houston is going up 0.0051 deg every year; in other words, **0.51 deg every 100 years**.

g) Similarly, the slope of the trend line for Global is 0.0073. This means that average global temperature is going up 0.0073 deg every year; in other words, **0.73 deg every 100 years**. **This indicates a slightly higher rate of global temperature increase as compared to Houston.**

h) The **correlation coefficient** between Houston and Global averages **is 0.59**, indicating that Houston's average temperature has a moderately positive impact on global average temperature.

K2 =CORREL(D2:D195,E2:E195)

	A	B	C	D	E	F	G	H	I	J	K
1	city	country	Ye	Housto	World	Houston-10YMA	World-10YMA	Diff	Mean	SD	Correlation
2	Houston	United States	1820	19.11	7.62			11.49	11.77	0.51	0.59
3	Houston	United States	1821	19.57	8.09			11.48			
4	Houston	United States	1822	20.05	8.19			11.86			
5	Houston	United States	1823	19.62	7.72			11.9			
6	Houston	United States	1824	20.19	8.55			11.64			

i) I added another city (Chicago) to the line chart to see if it followed the same trend as Houston and Global. Turns out, it did. Temperatures in Chicago are rising at the rate of 0.67 deg every 100 years. Chicago is much cooler than Houston but slightly warmer than global averages. The weather tends to fluctuate much more as displayed by the choppiness in the line chart.