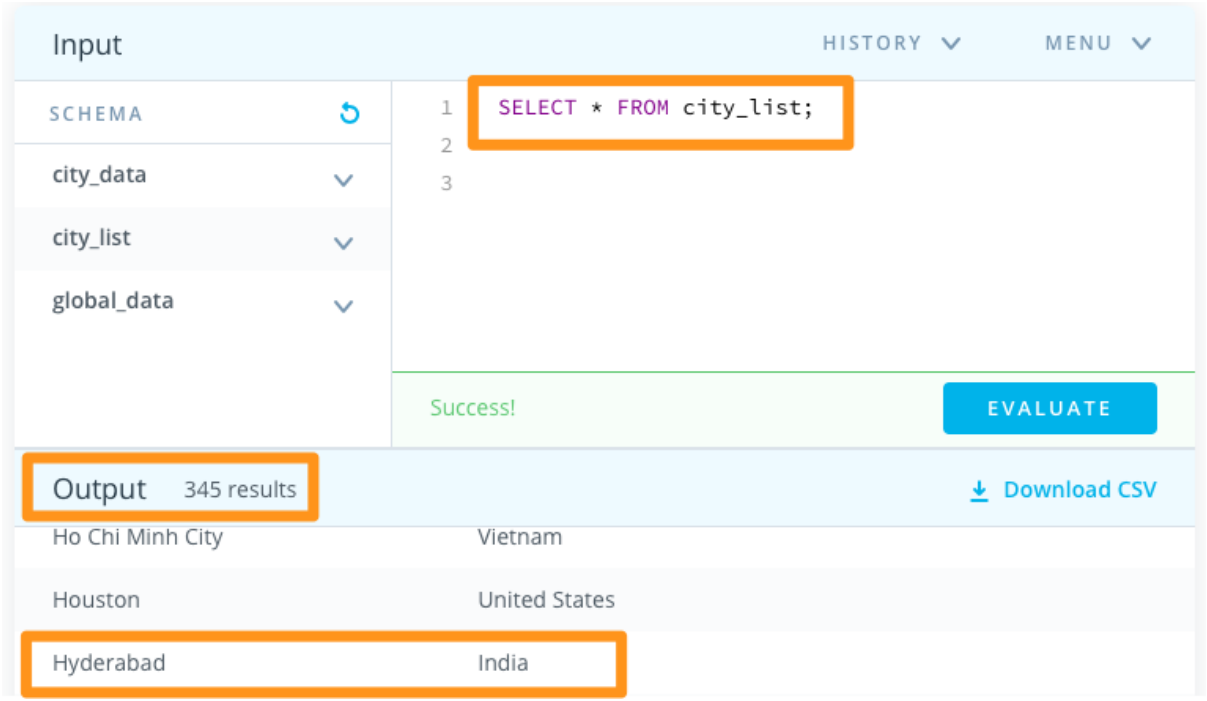
Project: Explore Weather Trends

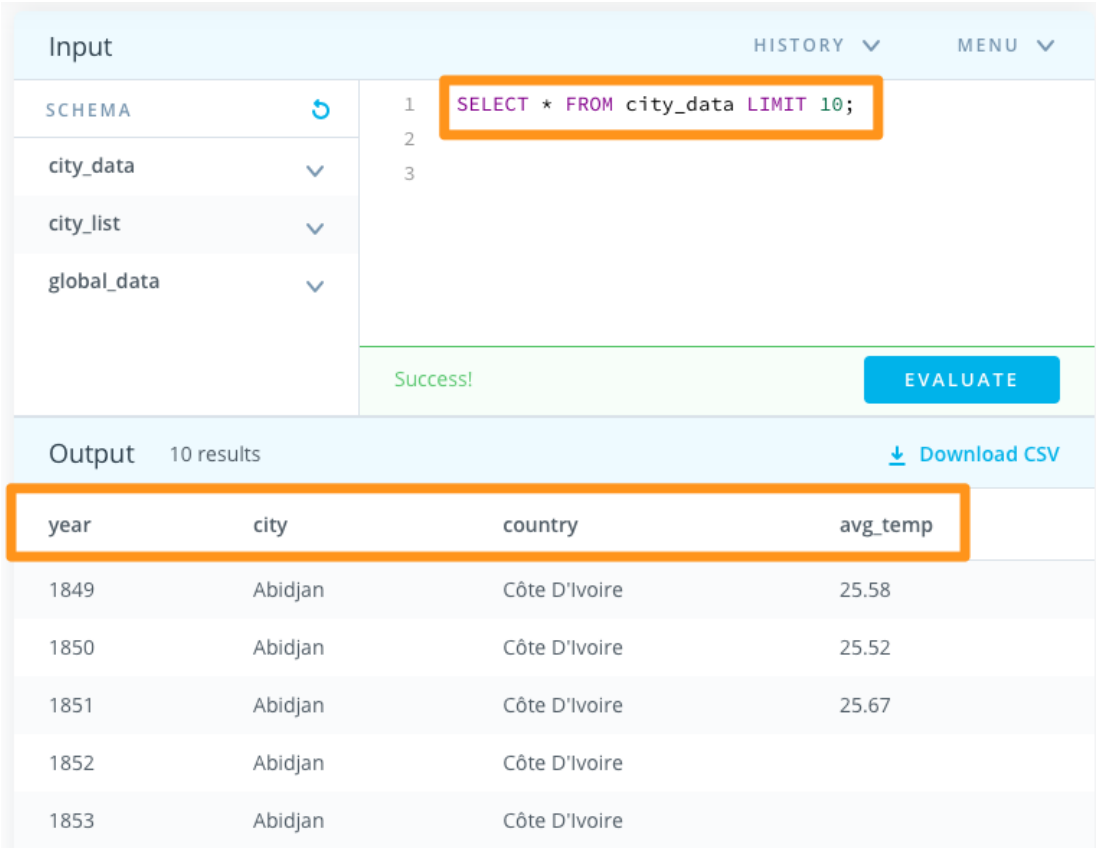
In this project we'll analyse local and global temperature data and compare the temperature trends to where I live, which is Hyderabad, India.

To do this, the following steps were followed:

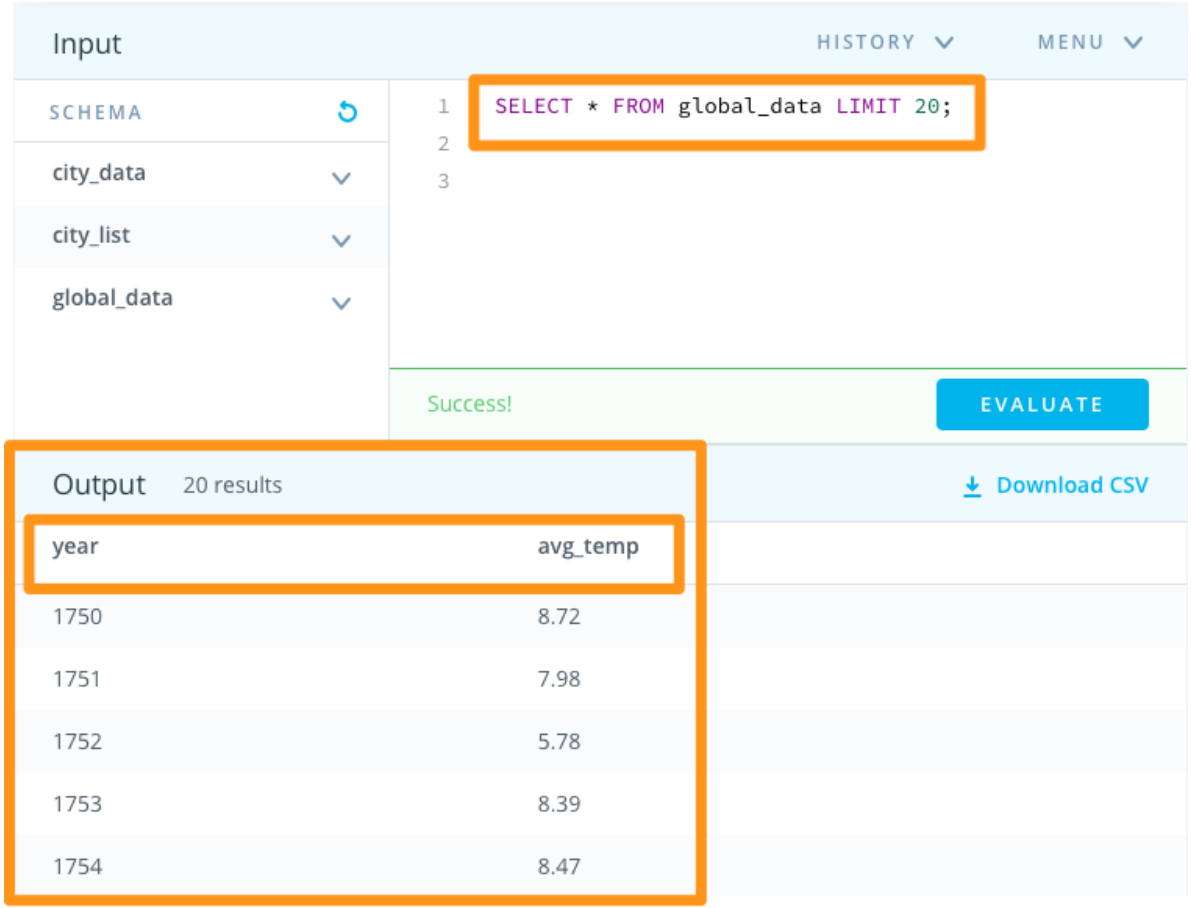
- **Extract the data** from the database.
    - Navigated to the workspace section which is connected to the database to extract data. This will be done using SQL.
    - There were three tables in the database:
        - city_list - This contained a list of cities and countries in the database. Used the following SQL query to find my city which is Hyderabad from India.
            - SELECT * FROM city_list;
            -



- city_data - This contained the average temperature for each city by year (degree C). Used the following SQL query for a quick review of the table:
  - SELECT * FROM city_data LIMIT 10;

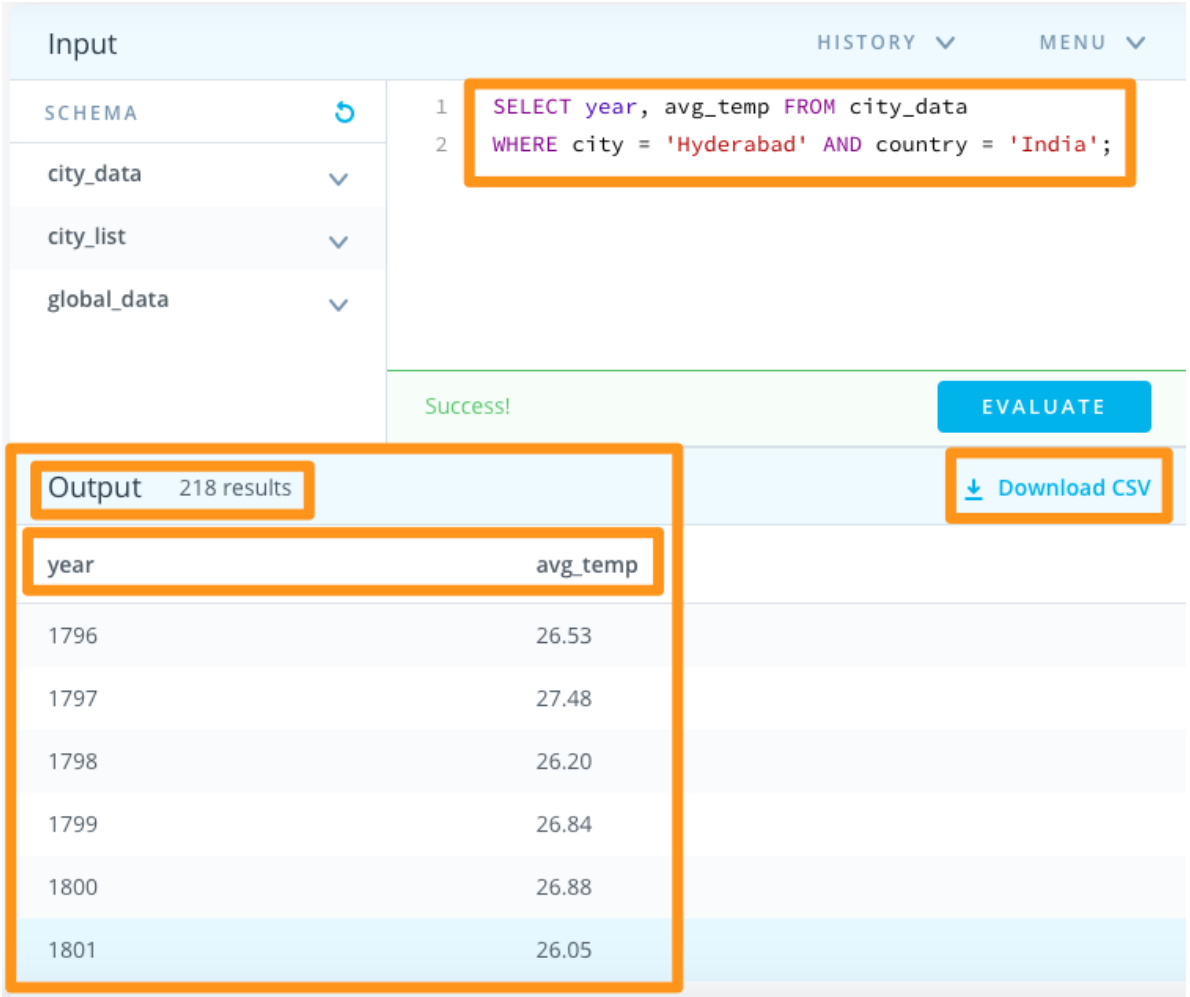


  - Observed that the table had the columns: year, city, country, avg_temp. Before extracting data from this table, I chose to have a quick view of the table global_data so as to know which columns might be required to create a line chart.
- global_data - This contains the average global temperatures by year (degree C). Use the following SQL query for a quick review of the table:
  - SELECT * FROM global_data LIMIT 20;



- Observed that the paper contained two columns: year and avg_temp. From this observation, I concluded that I need to extract the same two column data namely year and avg_temp from the table city_data for my city, which is Hyderabad, India.

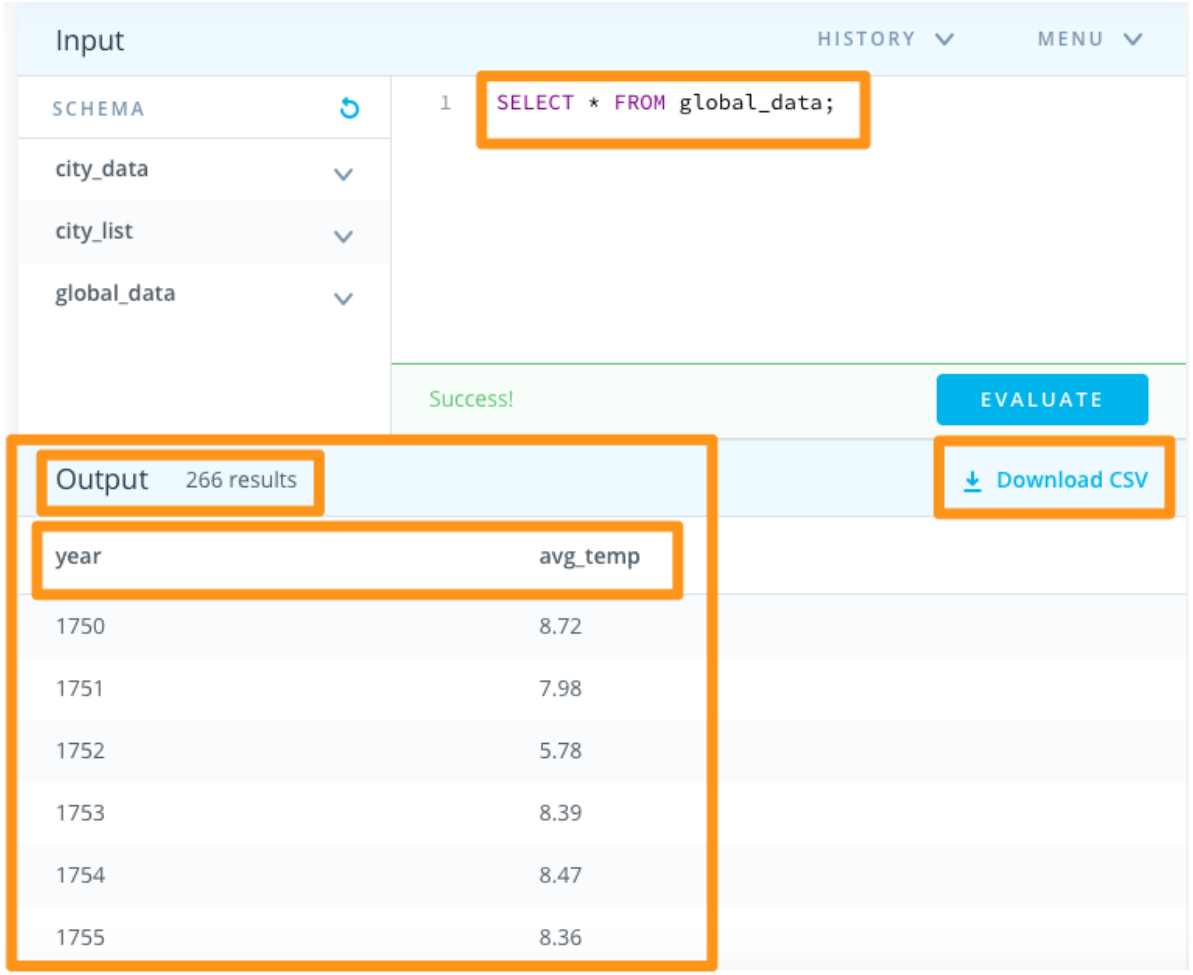
- **Queries for Extracting/Downloading the data.**
  - Used the following query to extract data namely year and avg_temp from the table city_data for my city, which is Hyderabad, India:
    - SELECT year, avg_temp FROM city_data WHERE city = 'Hyderabad' AND country = 'India';



- Output of the above query fetched 218 results with columns: year and avg_temp for city: Hyderabad, country: India.
- Clicked the “Download CSV” image URL to extract/download the data in a csv file on local drive.

▪ Used the following query to extract all data from the table global_data:

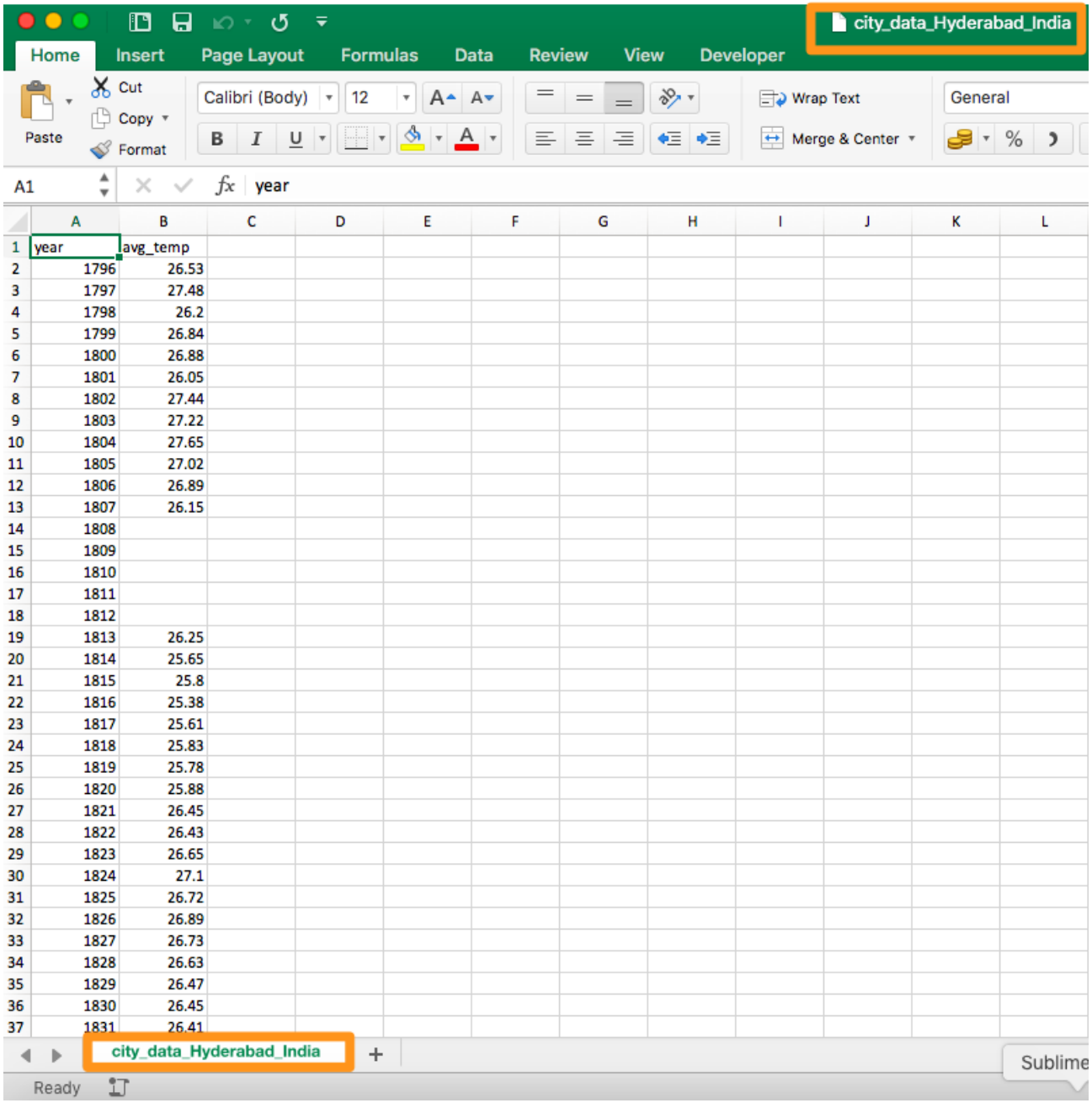
- SELECT * FROM global_data;



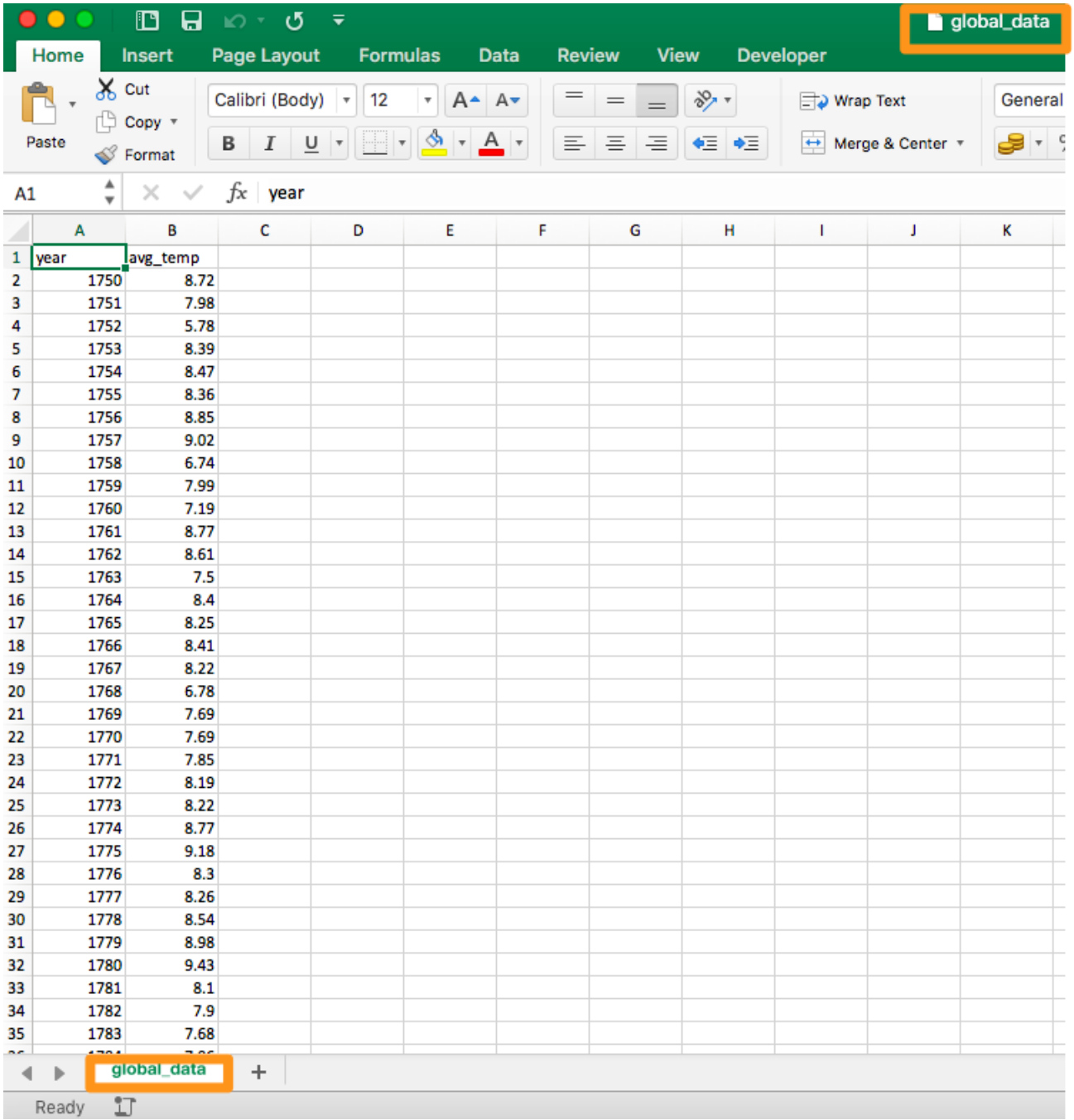
- Output of the above query fetched 266 results with columns: year and avg_temp.
- Clicked the “Download CSV” image URL to extract/download the data in a csv file on local drive.

  - This marks completion of the step of extraction of data.

- **Open up the CSV.**
  - Used Microsoft Excel to open up the csv files named “city_data_Hyderabad_India.csv” and “global_data.csv”.

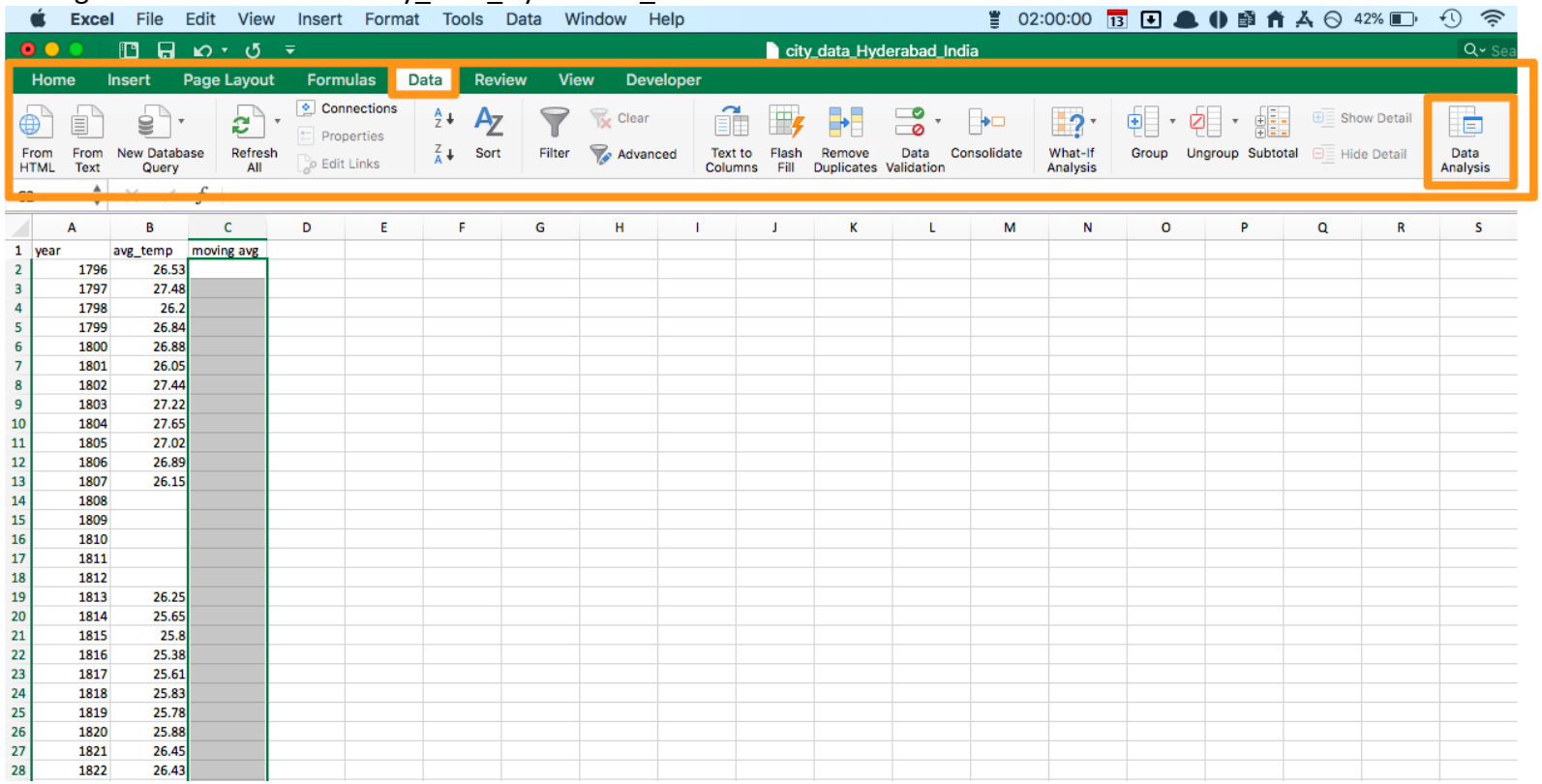


	A	B	C	D	E	F	G	H	I	J	K	L
1	year	avg_temp										
2	1796	26.53										
3	1797	27.48										
4	1798	26.2										
5	1799	26.84										
6	1800	26.88										
7	1801	26.05										
8	1802	27.44										
9	1803	27.22										
10	1804	27.65										
11	1805	27.02										
12	1806	26.89										
13	1807	26.15										
14	1808											
15	1809											
16	1810											
17	1811											
18	1812											
19	1813	26.25										
20	1814	25.65										
21	1815	25.8										
22	1816	25.38										
23	1817	25.61										
24	1818	25.83										
25	1819	25.78										
26	1820	25.88										
27	1821	26.45										
28	1822	26.43										
29	1823	26.65										
30	1824	27.1										
31	1825	26.72										
32	1826	26.89										
33	1827	26.73										
34	1828	26.63										
35	1829	26.47										
36	1830	26.45										
37	1831	26.41										



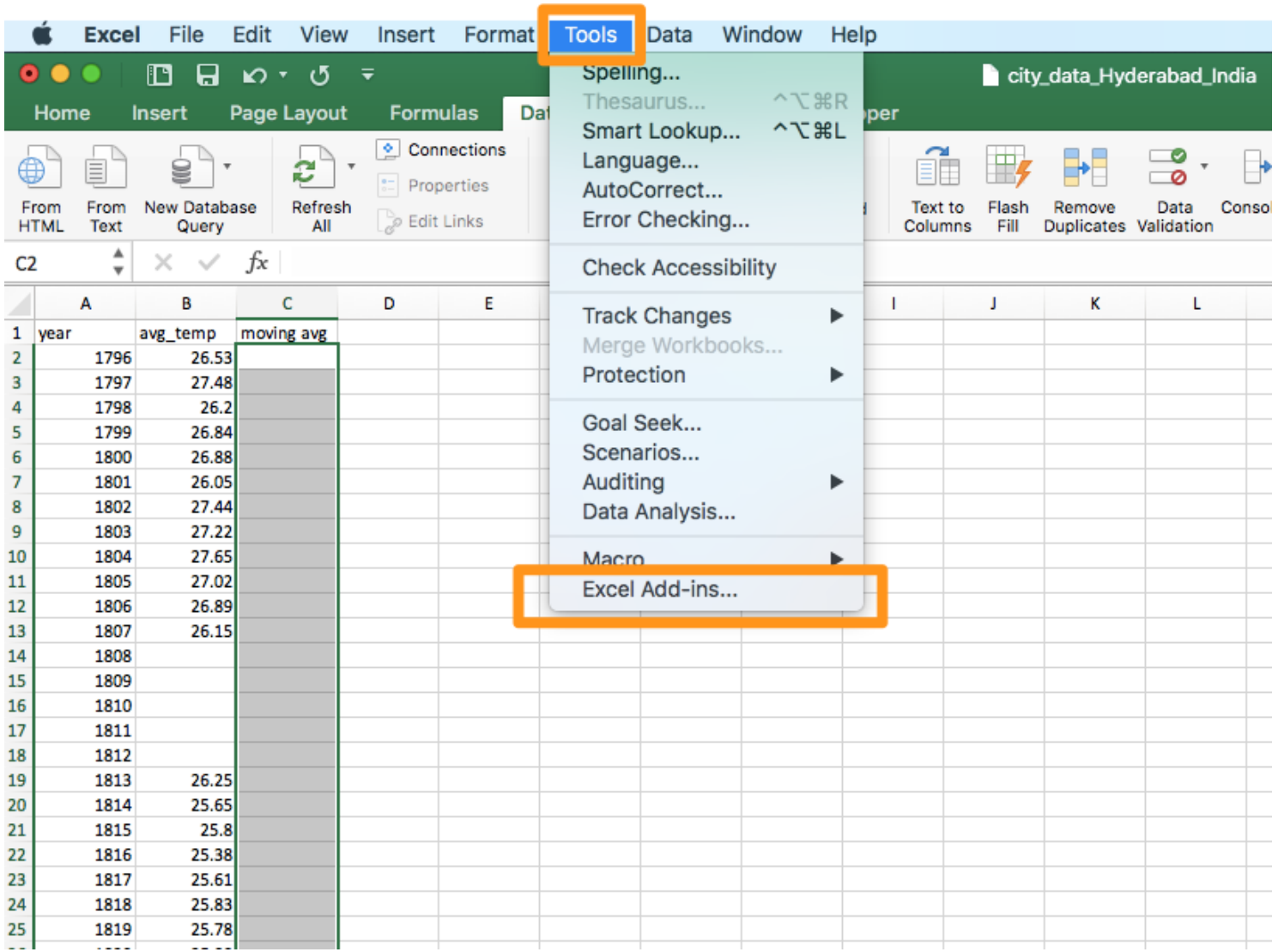
	year	avg_temp
1	year	avg_temp
2	1750	8.72
3	1751	7.98
4	1752	5.78
5	1753	8.39
6	1754	8.47
7	1755	8.36
8	1756	8.85
9	1757	9.02
10	1758	6.74
11	1759	7.99
12	1760	7.19
13	1761	8.77
14	1762	8.61
15	1763	7.5
16	1764	8.4
17	1765	8.25
18	1766	8.41
19	1767	8.22
20	1768	6.78
21	1769	7.69
22	1770	7.69
23	1771	7.85
24	1772	8.19
25	1773	8.22
26	1774	8.77
27	1775	9.18
28	1776	8.3
29	1777	8.26
30	1778	8.54
31	1779	8.98
32	1780	9.43
33	1781	8.1
34	1782	7.9
35	1783	7.68

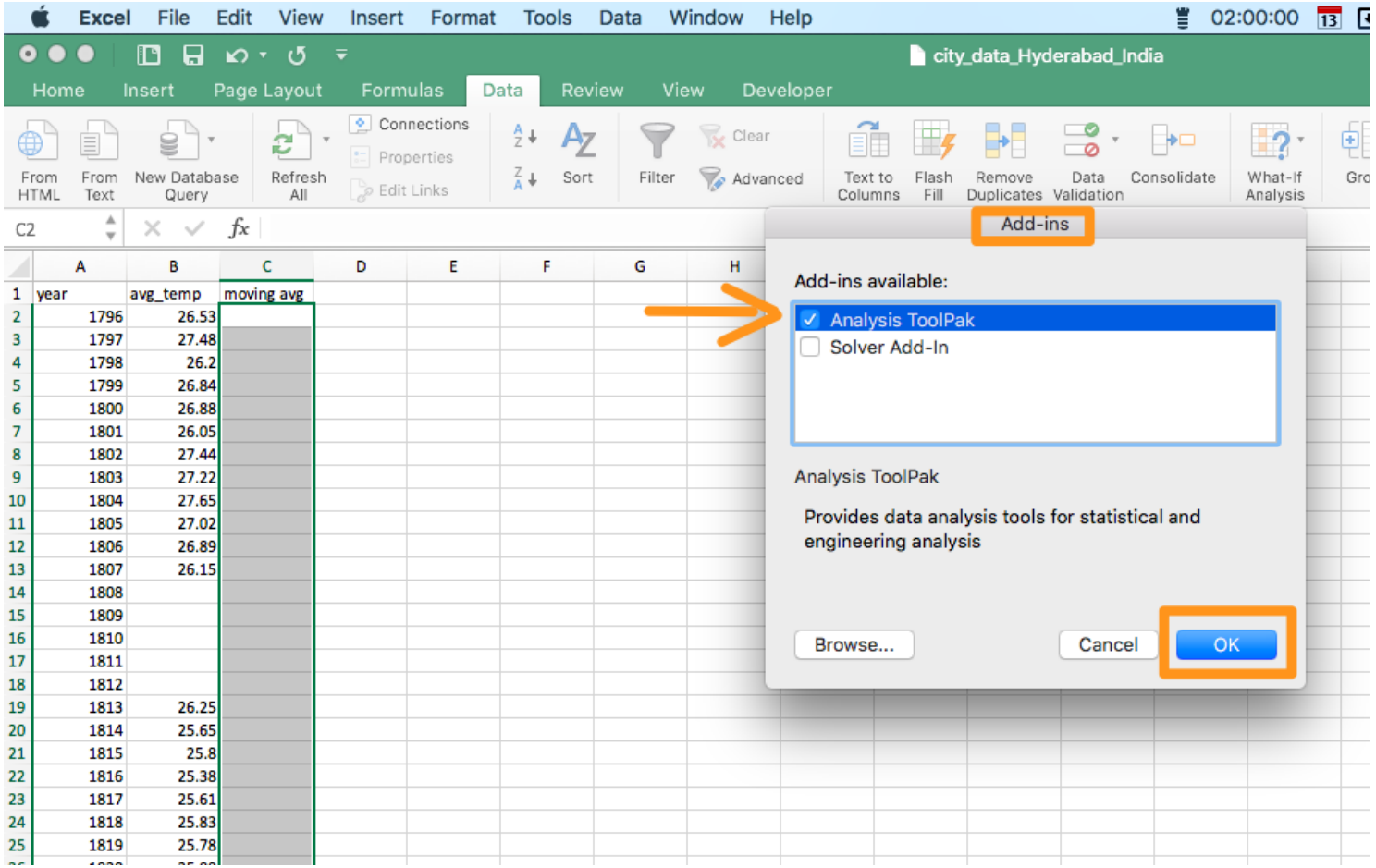
- **Create a line chart** that compares your temperature with the global temperatures.
    - **Calculate moving average**. Moving average is used for plotting instead of yearly averages in order to smooth out the lines and make trends more observable. Up
        - For the file “city_data_Hyderabad_India.csv”.
        - Check whether “Data Analysis” option is present by clicking “Data” tab from the ribbon and toolbar section of excel, which is being used to view the file “city_data_Hyderabad_India.csv”.



	A	B	C
1	year	avg_temp	moving avg
2	1796	26.53	
3	1797	27.48	
4	1798	26.2	
5	1799	26.84	
6	1800	26.88	
7	1801	26.05	
8	1802	27.44	
9	1803	27.22	
10	1804	27.65	
11	1805	27.02	
12	1806	26.89	
13	1807	26.15	
14	1808		
15	1809		
16	1810		
17	1811		
18	1812		
19	1813	26.25	
20	1814	25.65	
21	1815	25.8	
22	1816	25.38	
23	1817	25.61	
24	1818	25.83	
25	1819	25.78	
26	1820	25.88	
27	1821	26.45	
28	1822	26.43	

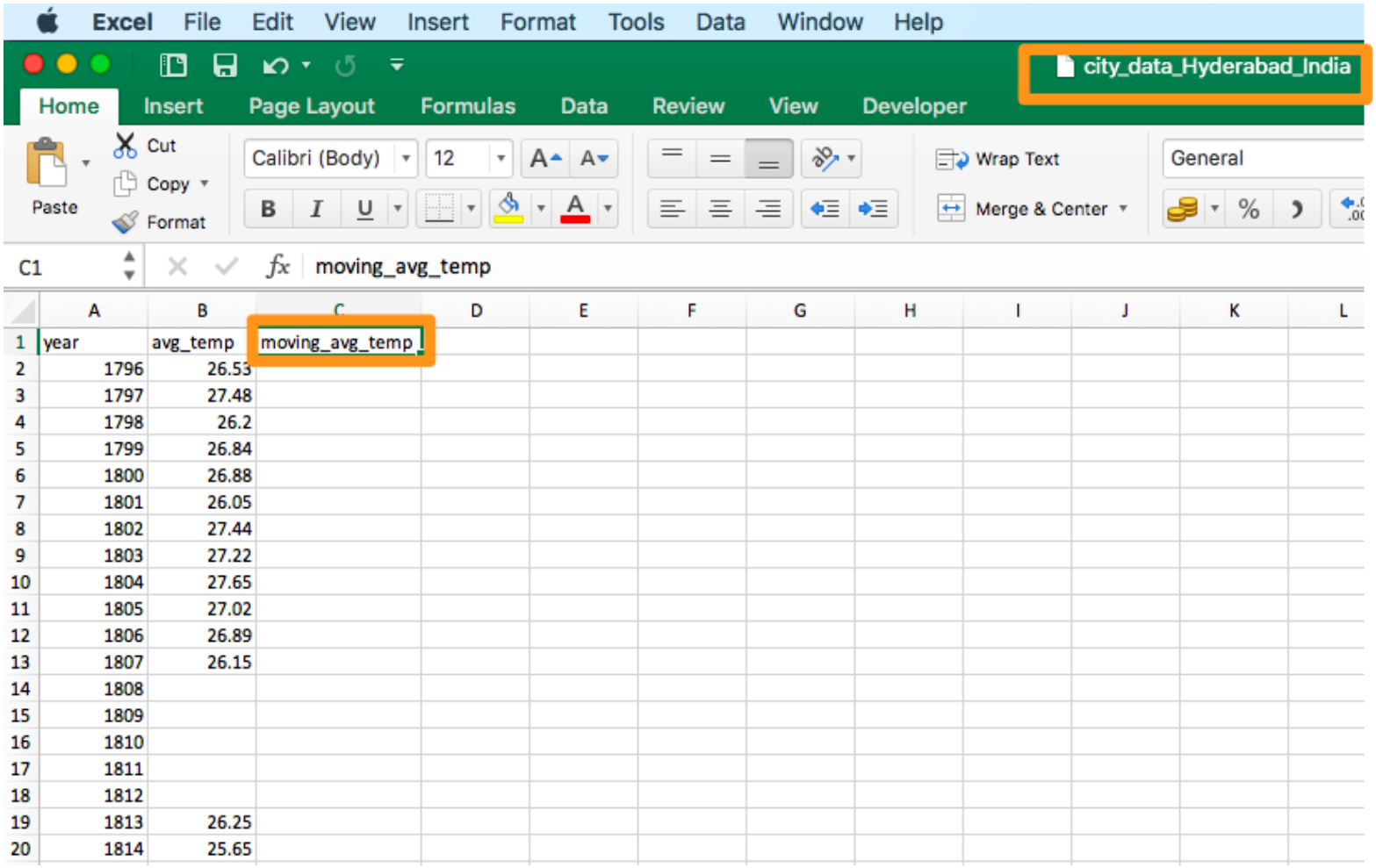
- If not, click Tools -> Excel Add-ins… from the File menu. Add-ins dialog box opens up. Select the checkbox for option “Analysis ToolPak” and click “OK” button. After this process one should see “Data Analysis” option by clicking “Data” tab under the ribbons and toobars section of the excel or “Tools” menu.


Excel File Edit View Insert Format Tools Data Window Help
city_data_Hyderabad_India
Home Insert Page Layout Formulas
From HTML
From Text
New Database Query
Refresh All
Connections
Properties
Edit Links
Text to Columns
Flash Fill
Remove Duplicates
Data Validation
C2
Spelling...
Thesaurus... ^⌥⌘R
Smart Lookup... ^⌥⌘L
Language...
AutoCorrect...
Error Checking...
Check Accessibility
Track Changes
Merge Workbooks...
Protection
Goal Seek...
Scenarios...
Auditing
Data Analysis...
Macro
Excel Add-ins...
year avg_temp moving avg
1796 26.53
1797 27.48
1798 26.2
1799 26.84
1800 26.88
1801 26.05
1802 27.44
1803 27.22
1804 27.65
1805 27.02
1806 26.89
1807 26.15
1808
1809
1810
1811
1812
1813 26.25
1814 25.65
1815 25.8
1816 25.38
1817 25.61
1818 25.83
1819 25.78



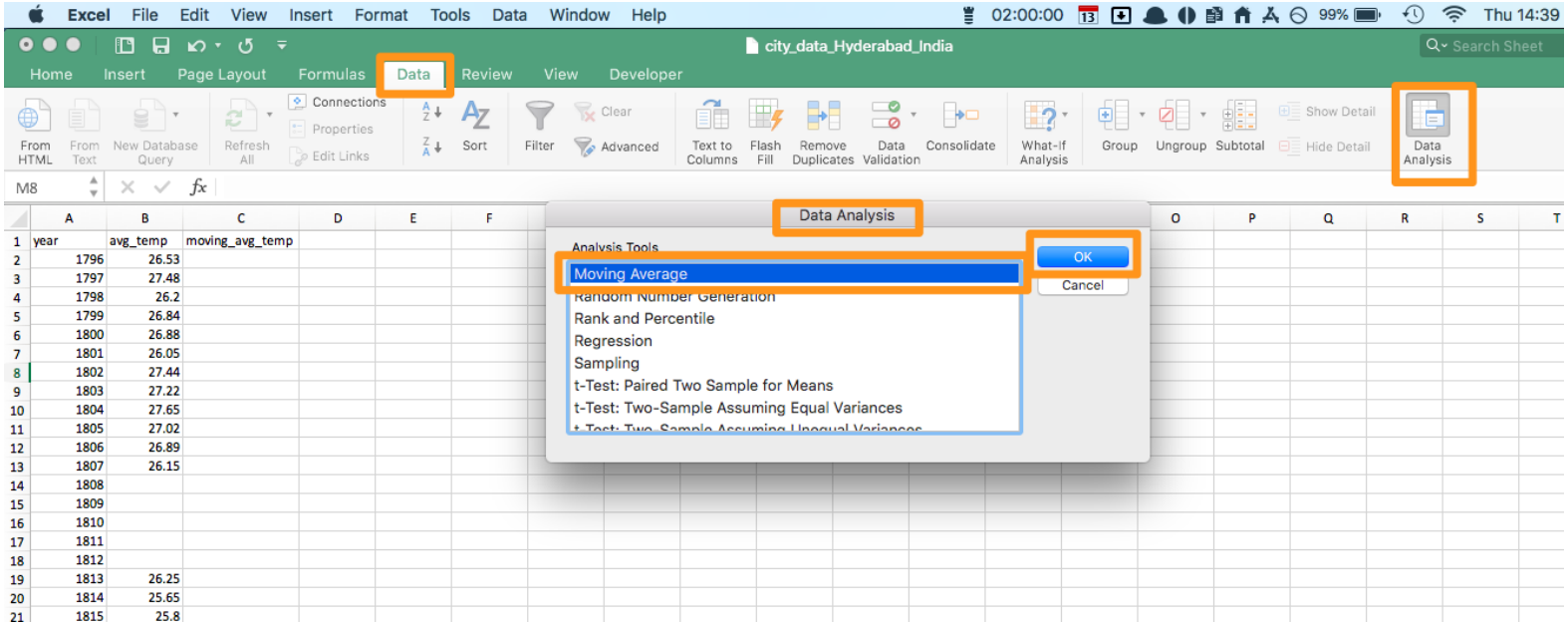
	A	B	C
1	year	avg_temp	moving avg
2	1796	26.53	
3	1797	27.48	
4	1798	26.2	
5	1799	26.84	
6	1800	26.88	
7	1801	26.05	
8	1802	27.44	
9	1803	27.22	
10	1804	27.65	
11	1805	27.02	
12	1806	26.89	
13	1807	26.15	
14	1808		
15	1809		
16	1810		
17	1811		
18	1812		
19	1813	26.25	
20	1814	25.65	
21	1815	25.8	
22	1816	25.38	
23	1817	25.61	
24	1818	25.83	
25	1819	25.78	

- Added another column label “moving_avg_temp” in file “city_data_Hyderabad_India.csv”.

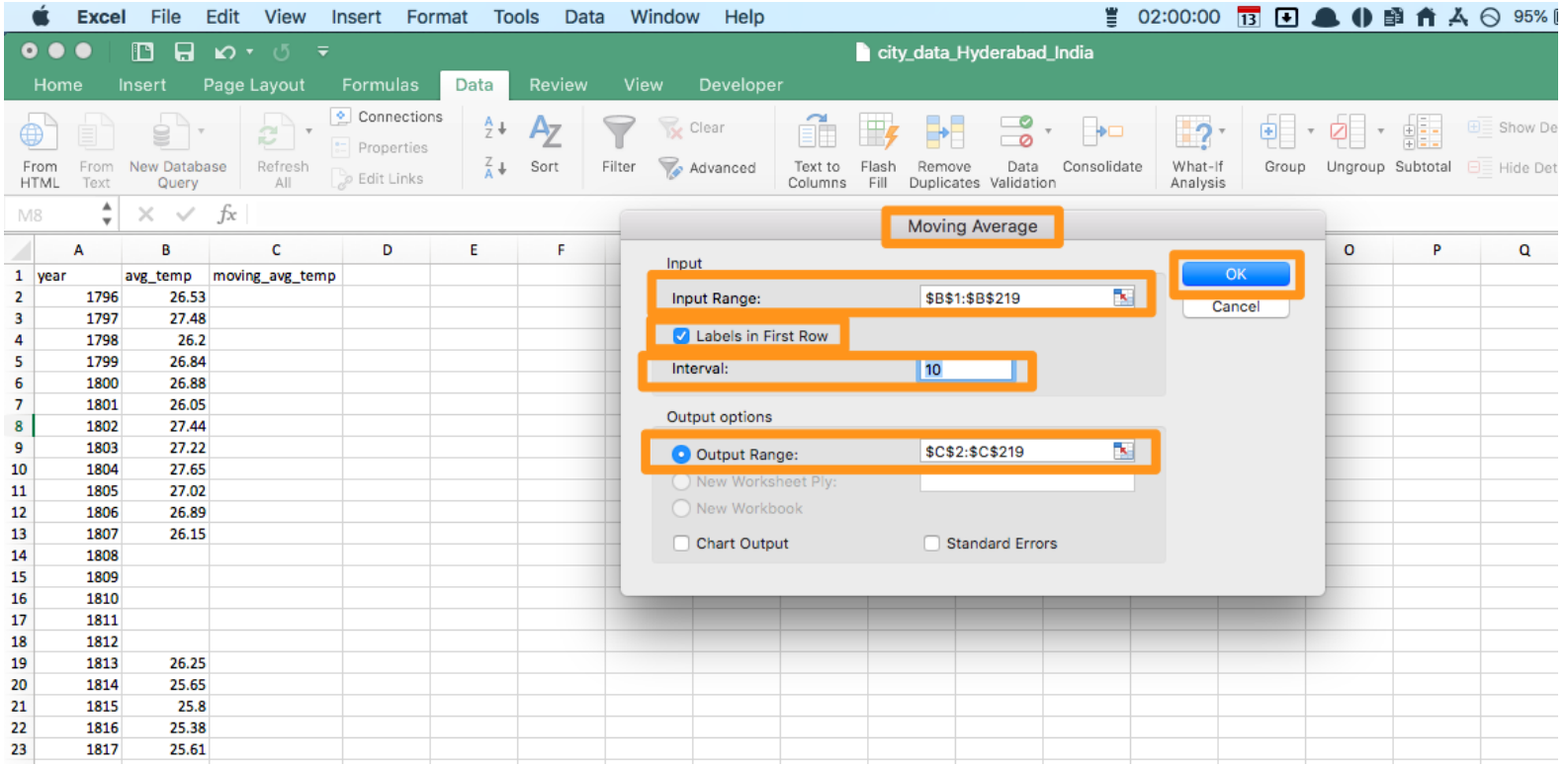


	A	B	C
1	year	avg_temp	moving_avg_temp
2	1796	26.53	
3	1797	27.48	
4	1798	26.2	
5	1799	26.84	
6	1800	26.88	
7	1801	26.05	
8	1802	27.44	
9	1803	27.22	
10	1804	27.65	
11	1805	27.02	
12	1806	26.89	
13	1807	26.15	
14	1808		
15	1809		
16	1810		
17	1811		
18	1812		
19	1813	26.25	
20	1814	25.65	

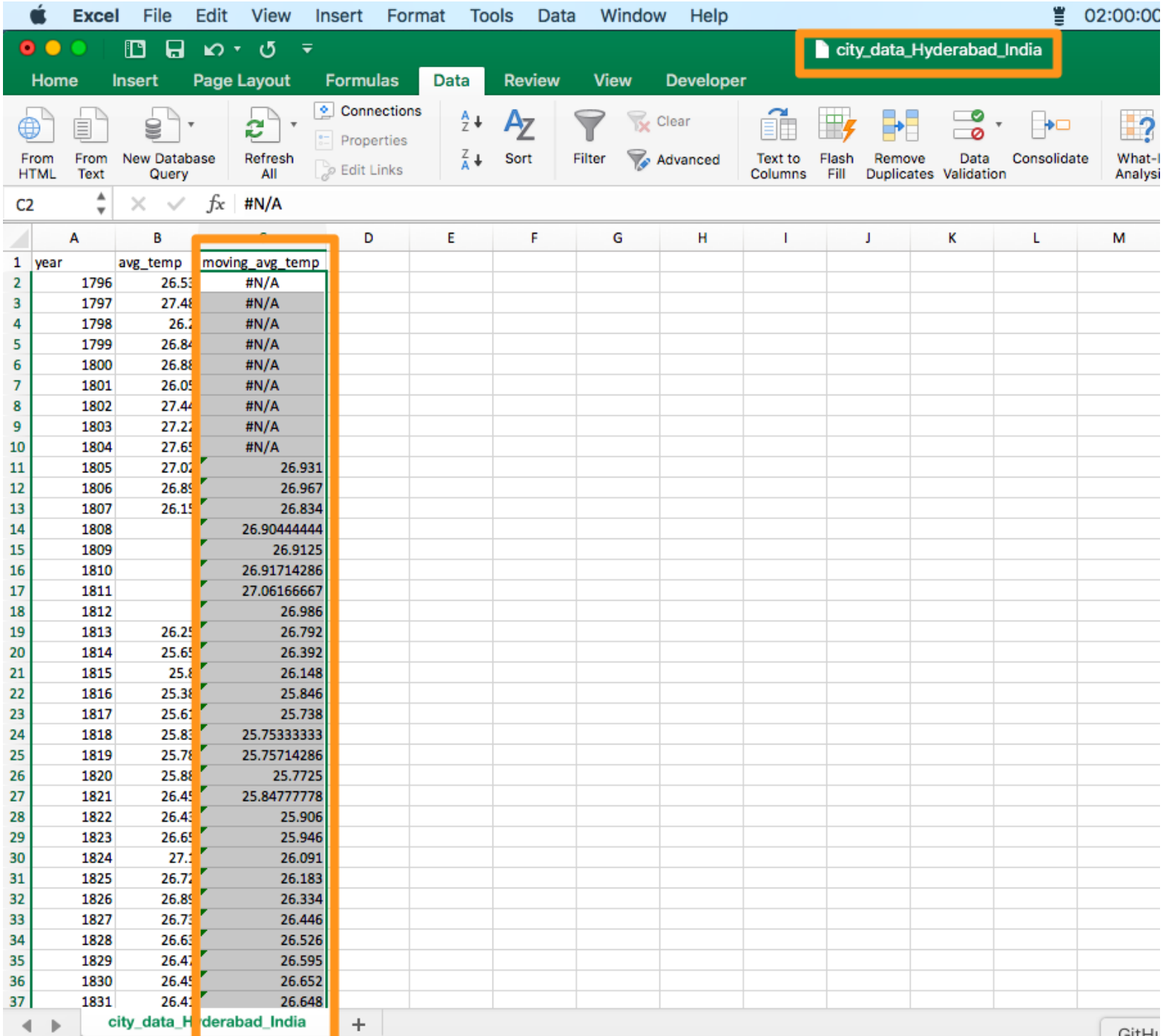
- Clicked Data – Data Analysis from the ribbon and toolbars section of the excel to obtain the “Data Analysis” dialog box. Selected “Moving Average” option from under “Analysis Tools” and clicked “OK” button.



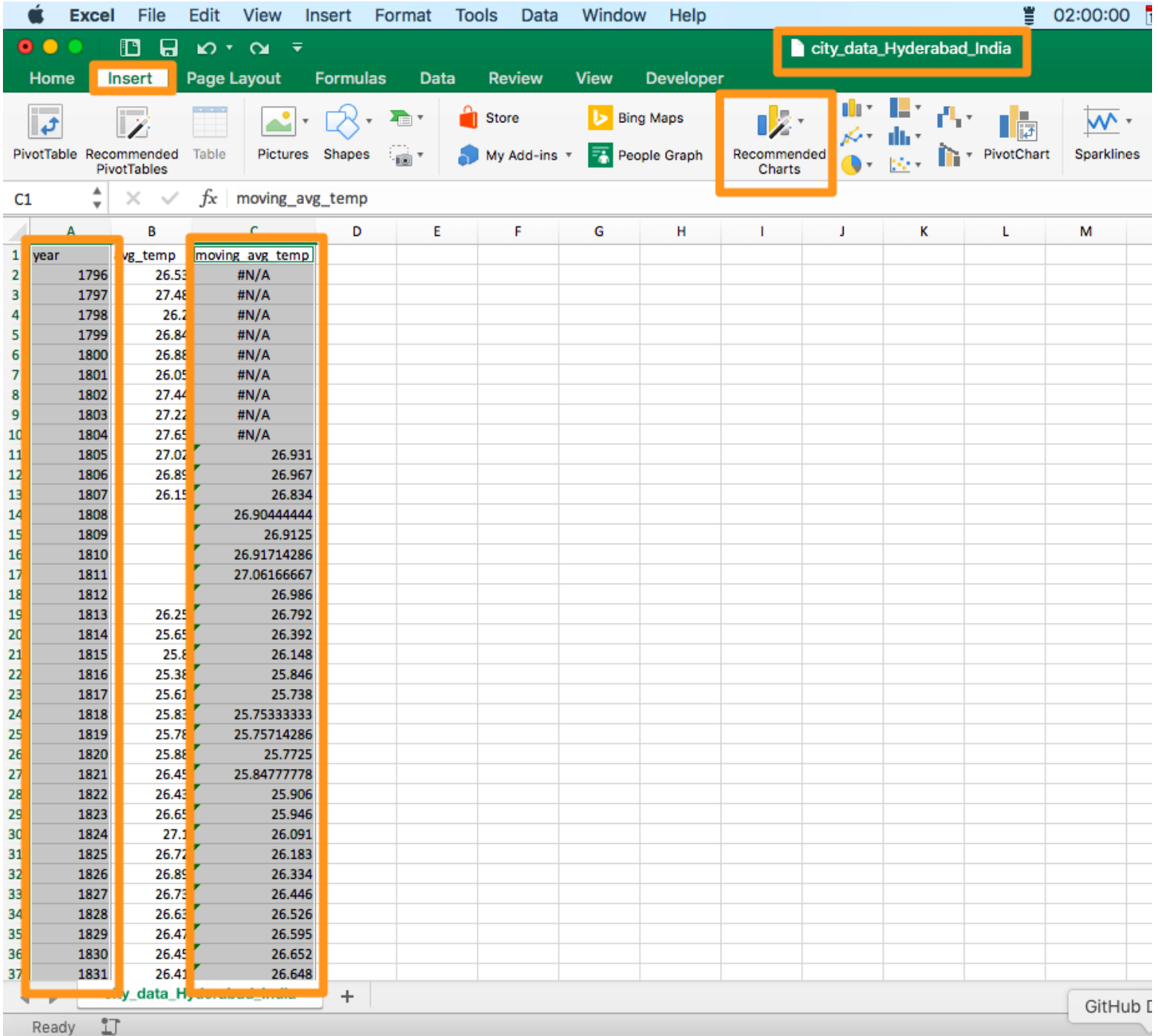
- User is transferred to the “Moving Average” dialog box. Selected the “Input Range” by selecting all the data cells in the column “avg_temp” including the column header and checked the checkbox “Label in First Row”. Entered 10 as the value for the text box input “Interval”. Similarly, selected the “Output Range” by selecting all the blank cells in the column “moving_avg_temp” excluding the column header. Clicked the “OK” button to find that the moving average of temperature values get filled in the “moving_avg_temp” column of the file “city_data_Hyderabad_India.csv” in excel.

- 

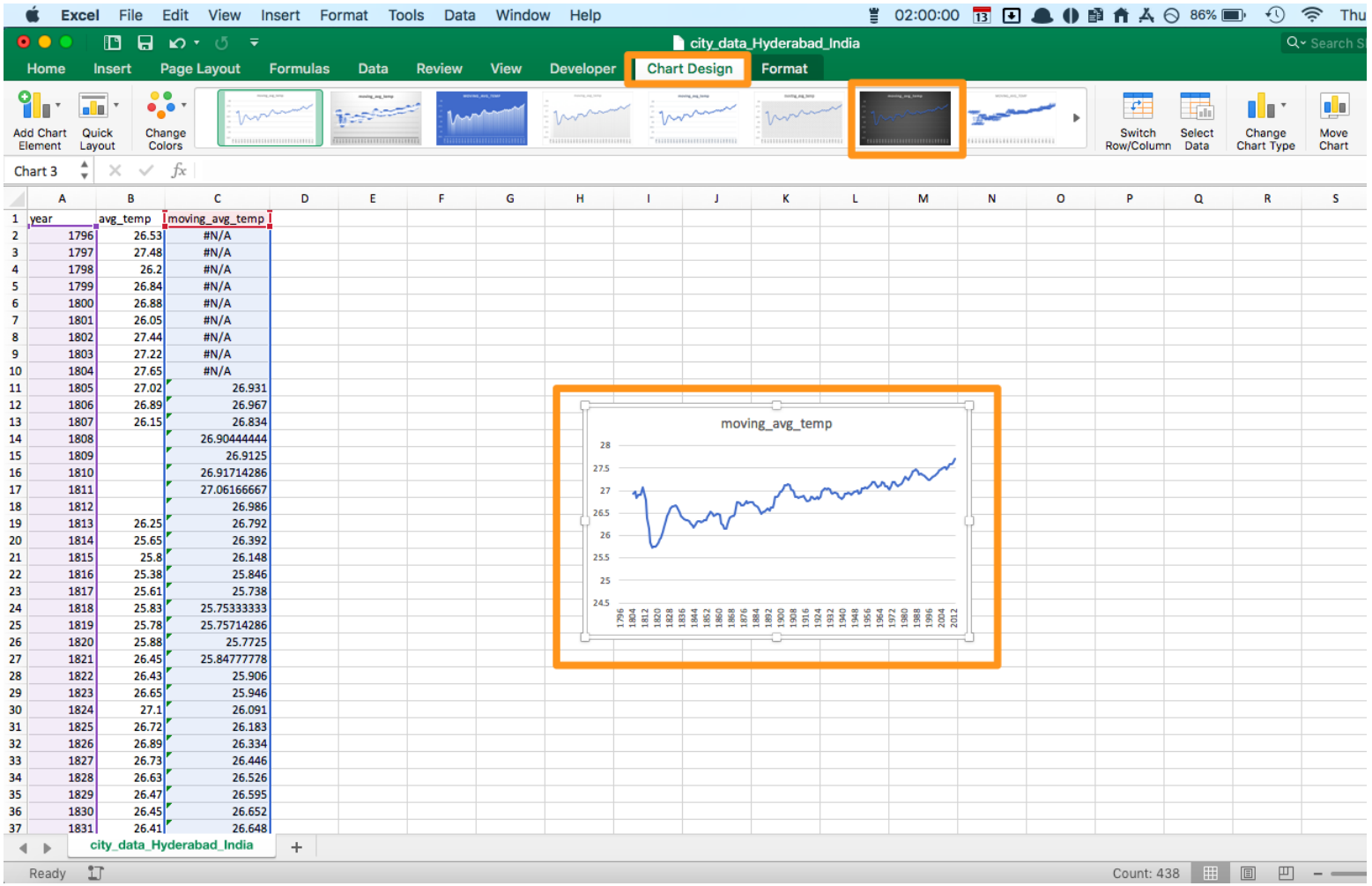
	A	B	C
1	year	avg_temp	moving_avg_temp
2	1796	26.53	
3	1797	27.48	
4	1798	26.2	
5	1799	26.84	
6	1800	26.88	
7	1801	26.05	
8	1802	27.44	
9	1803	27.22	
10	1804	27.65	
11	1805	27.02	
12	1806	26.89	
13	1807	26.15	
14	1808		
15	1809		
16	1810		
17	1811		
18	1812		
19	1813	26.25	
20	1814	25.65	
21	1815	25.8	
22	1816	25.38	
23	1817	25.61	


Excel File Edit View Insert Format Tools Data Window Help
02:00:00
city_data_Hyderabad_India
Home Insert Page Layout Formulas Data Review View Developer
From HTML
From Text
New Database Query
Refresh All
Connections
Properties
Edit Links
Sort
Filter
Clear
Advanced
Text to Columns
Flash Fill
Remove Duplicates
Data Validation
Consolidate
C2
#N/A
year
avg_temp
moving_avg_temp
1796 #N/A
1797 #N/A
1798 #N/A
1799 #N/A
1800 #N/A
1801 #N/A
1802 #N/A
1803 #N/A
1804 #N/A
1805 26.931
1806 26.967
1807 26.834
1808 26.90444444
1809 26.9125
1810 26.91714286
1811 27.06166667
1812 26.986
1813 26.792
1814 26.392
1815 26.148
1816 25.846
1817 25.738
1818 25.75333333
1819 25.75714286
1820 25.7725
1821 25.84777778
1822 25.906
1823 25.946
1824 26.091
1825 26.183
1826 26.334
1827 26.446
1828 26.526
1829 26.595
1830 26.652
1831 26.648
city_data_Hyderabad_India

        - Observed that the first 9 values filled as “#N/A” as the interval chosen by me was 10 and the data until the 9th row was insufficient to calculate the moving average temperature.
- **Create line chart.**
    - Selected both the columns namely “year” and “moving_avg_temp” from the file “city_data_Hyderabad_India.csv” in excel. Navigated to the “Insert” tab under the ribbons and toolbars section of the excel. Clicked the “Recommended Charts” option.

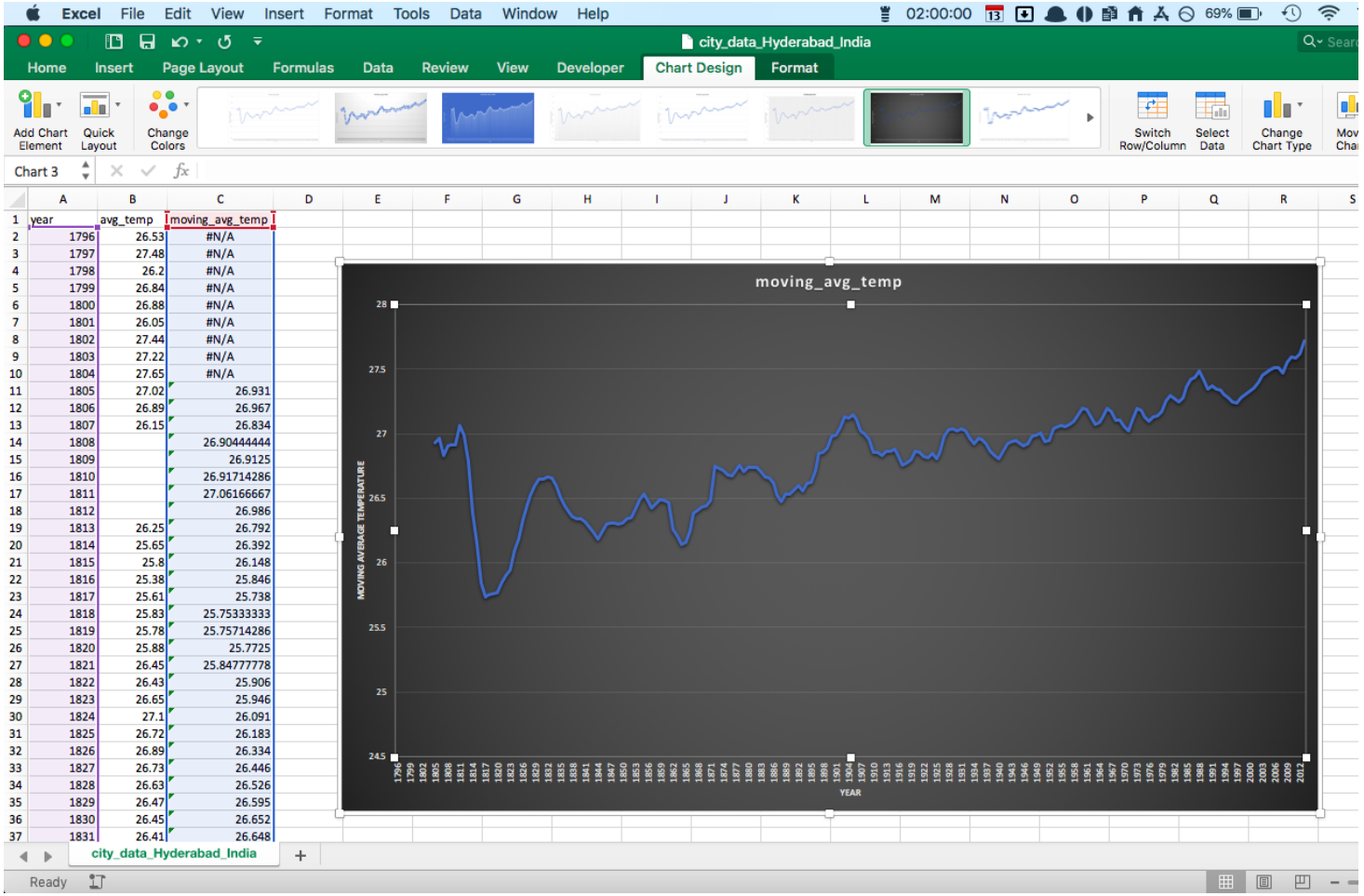

Excel File Edit View Insert Format Tools Data Window Help
02:00:00
city_data_Hyderabad_India
Home Insert Page Layout Formulas Data Review View Developer
PivotTable
Recommended PivotTables
Table
Pictures
Shapes
Store
My Add-ins
Bing Maps
People Graph
Recommended Charts
PivotChart
Sparklines
C1
moving_avg_temp
year
moving_avg_temp
#N/A
26.931
26.967
26.834
26.90444444
26.9125
26.91714286
27.06166667
26.986
26.792
26.392
26.148
25.846
25.738
25.75333333
25.75714286
25.7725
25.84777778
25.906
25.946
26.091
26.183
26.334
26.446
26.526
26.595
26.652
26.648
Ready
GitHub

- Selected Line chart option from the “Recommended Charts” drop down presented. Observed the line chart gets displayed in the excel.



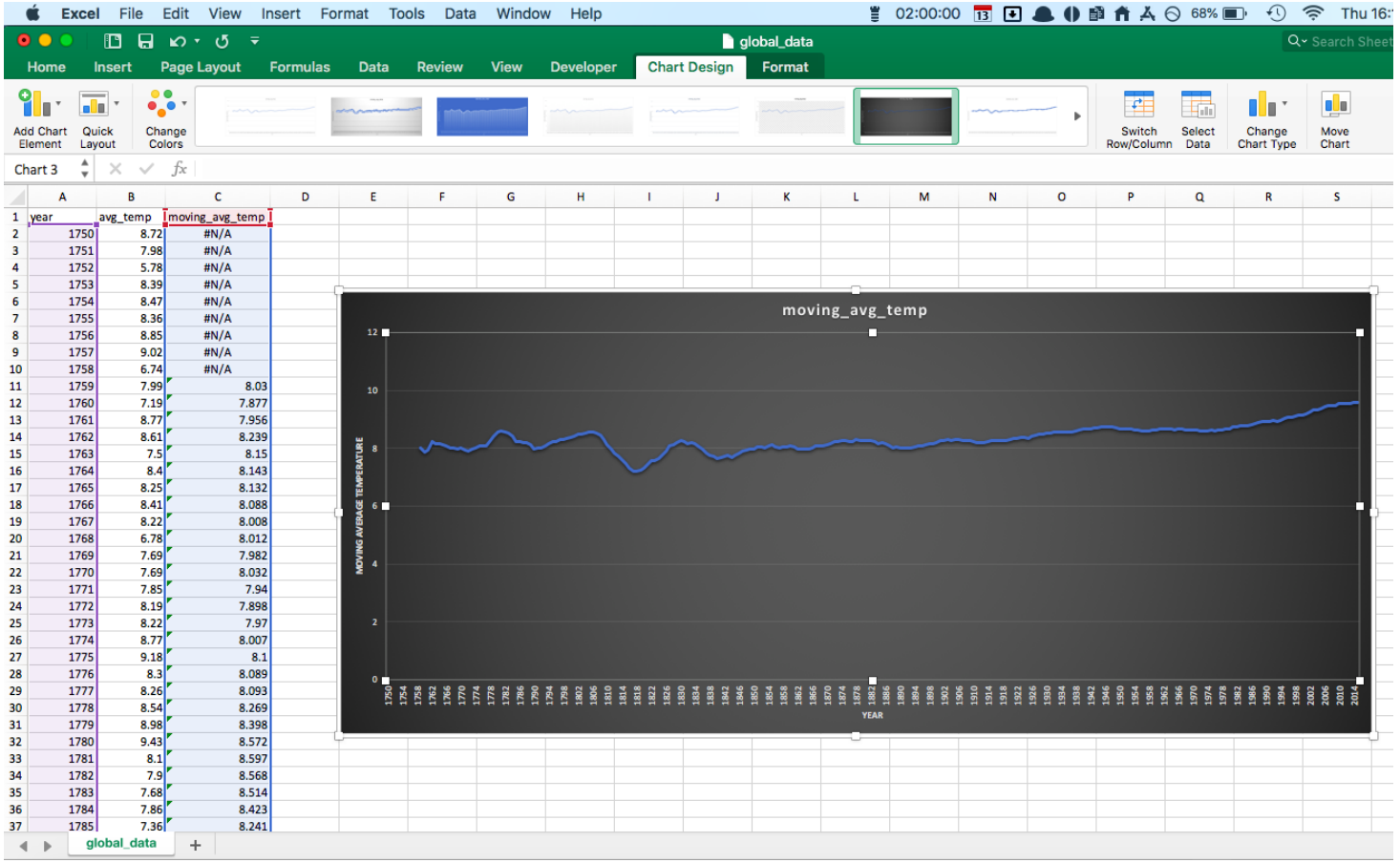
year	avg_temp	moving_avg_temp
1796	26.53	#N/A
1797	27.48	#N/A
1798	26.2	#N/A
1799	26.84	#N/A
1800	26.88	#N/A
1801	26.05	#N/A
1802	27.44	#N/A
1803	27.22	#N/A
1804	27.65	#N/A
1805	27.02	26.931
1806	26.89	26.967
1807	26.15	26.834
1808		26.90444444
1809		26.9125
1810		26.91714286
1811		27.06166667
1812		26.986
1813	26.25	26.792
1814	25.65	26.392
1815	25.8	26.148
1816	25.38	25.846
1817	25.61	25.738
1818	25.83	25.75333333
1819	25.78	25.75714286
1820	25.88	25.7725
1821	26.45	25.84777778
1822	26.43	25.906
1823	26.65	25.946
1824	27.1	26.091
1825	26.72	26.183
1826	26.89	26.334
1827	26.73	26.446
1828	26.63	26.526
1829	26.47	26.595
1830	26.45	26.652
1831	26.41	26.648

- Resized and moved the line chart and applied a dark chart design for better visualization.



year	avg_temp	moving_avg_temp
1796	26.53	#N/A
1797	27.48	#N/A
1798	26.2	#N/A
1799	26.84	#N/A
1800	26.88	#N/A
1801	26.05	#N/A
1802	27.44	#N/A
1803	27.22	#N/A
1804	27.65	#N/A
1805	27.02	26.931
1806	26.89	26.967
1807	26.15	26.834
1808		26.90444444
1809		26.9125
1810		26.91714286
1811		27.06166667
1812		26.986
1813	26.25	26.792
1814	25.65	26.392
1815	25.8	26.148
1816	25.38	25.846
1817	25.61	25.738
1818	25.83	25.75333333
1819	25.78	25.75714286
1820	25.88	25.7725
1821	26.45	25.84777778
1822	26.43	25.906
1823	26.65	25.946
1824	27.1	26.091
1825	26.72	26.183
1826	26.89	26.334
1827	26.73	26.446
1828	26.63	26.526
1829	26.47	26.595
1830	26.45	26.652
1831	26.41	26.648

- For the file “global_data.csv”.
    - Followed the similar steps to draw a line chart as described above (calculated moving average temperature with 10-year interval as done above), and obtained the line chart for the file “global_data.csv” as displayed below.

- 

year	avg_temp	moving_avg_temp
1750	8.72	#N/A
1751	7.98	#N/A
1752	5.78	#N/A
1753	8.39	#N/A
1754	8.47	#N/A
1755	8.36	#N/A
1756	8.85	#N/A
1757	9.02	#N/A
1758	6.74	#N/A
1759	7.99	8.03
1760	7.19	7.877
1761	8.77	7.956
1762	8.61	8.239
1763	7.5	8.15
1764	8.4	8.143
1765	8.25	8.132
1766	8.41	8.088
1767	8.22	8.008
1768	6.78	8.012
1769	7.69	7.982
1770	7.69	8.032
1771	7.85	7.94
1772	8.19	7.898
1773	8.22	7.97
1774	8.77	8.007
1775	9.18	8.1
1776	8.3	8.089
1777	8.26	8.093
1778	8.54	8.269
1779	8.98	8.398
1780	9.43	8.572
1781	8.1	8.597
1782	7.9	8.568
1783	7.68	8.514
1784	7.86	8.423
1785	7.36	8.241

- **Observations.**
  - It is found that the city of Hyderabad from India seems hotter on average compared to global average temperature.
  - Data shows that average temperature for the world has been steadily increasing over the years. The average temperature for Hyderabad from India also shows similar increasing pattern.
  - As per the data from the line chart for global average temperature, the years 1814 to 1820 saw a marked decrease in the average world temperature. This can also be verified from the line chart for the city of Hyderabad from India.

- Other than the period mentioned above, the world temperature has been rising steadily over the years which drives us to the conclusion that the world is getting hotter. The trends in the world average temperature are also well reflected from the local data from the city of Hyderabad, India.