Data Analysis Nanodegree

Project:

Weather Trends

By

Taj Banook Bugti

Summary:

In this project, we are required to compare and analyze global temperature with our nearest city. So, I selected Karachi and compared the data with global temperature's data.

Extracting data (first step):

I pulled out data from database schema in workspace using following SQL query:

```
SELECT city, gd.year,gd.avg_temp AS gd_temp, cd.avg_temp AS cd_temp

FROM city_data AS cd

FULL JOIN global_data AS gd

ON cd.year=gd.year
```

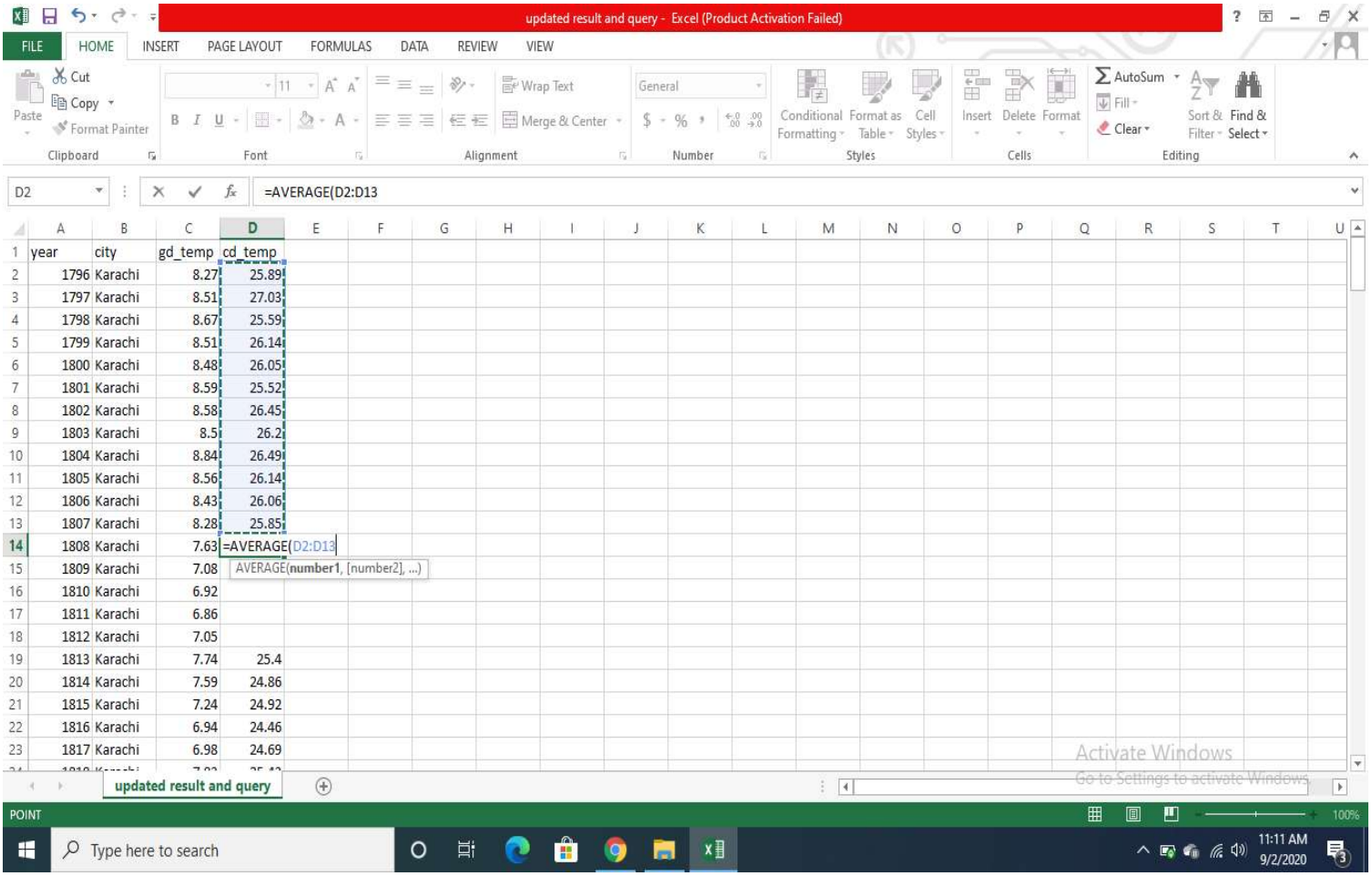
Importing data (second step):

I imported the results to Excel worksheet.

Handling missing data(third step):

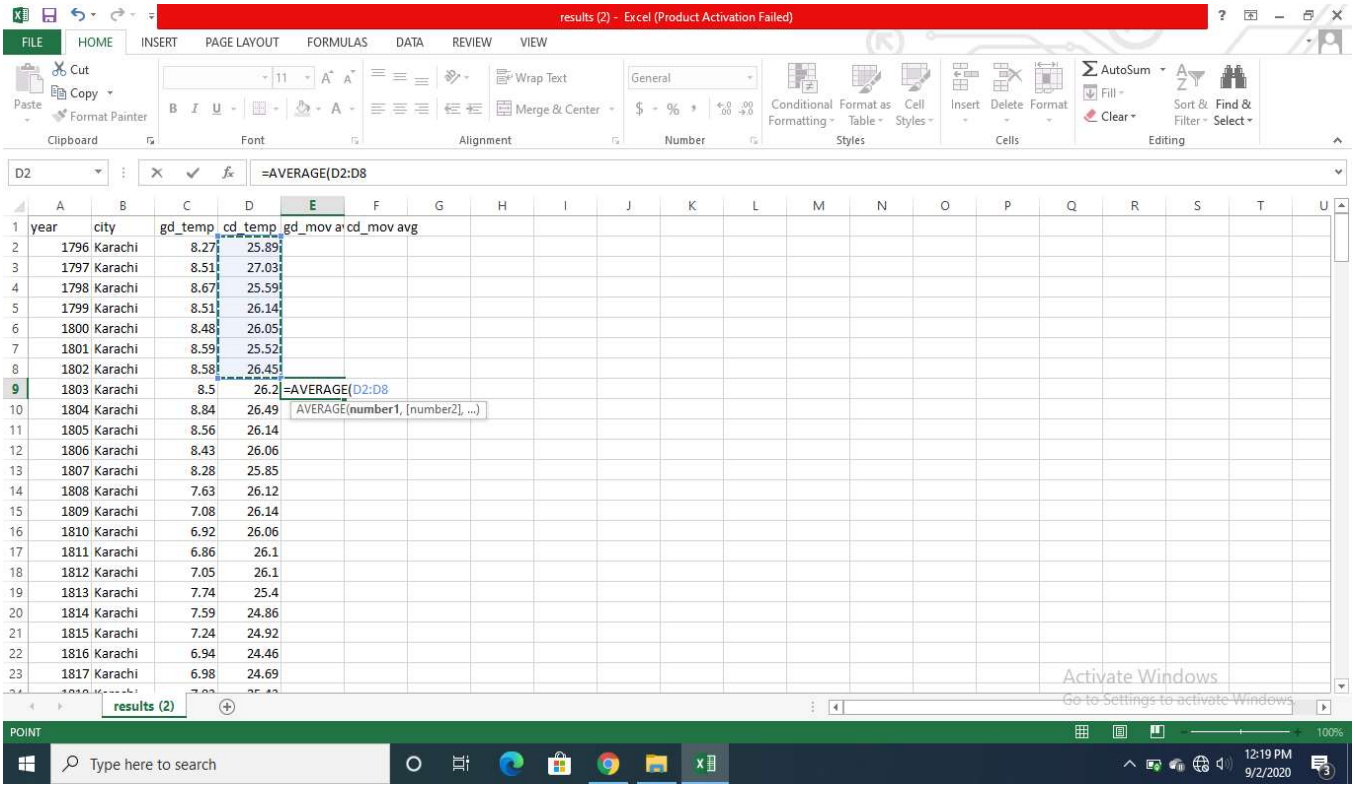
I calculated average of previous years and inserted the value in empty cells. I also added obtained results (averages) from above cells, recalculated like in moving average and inserted those values in below cells (empty cells).

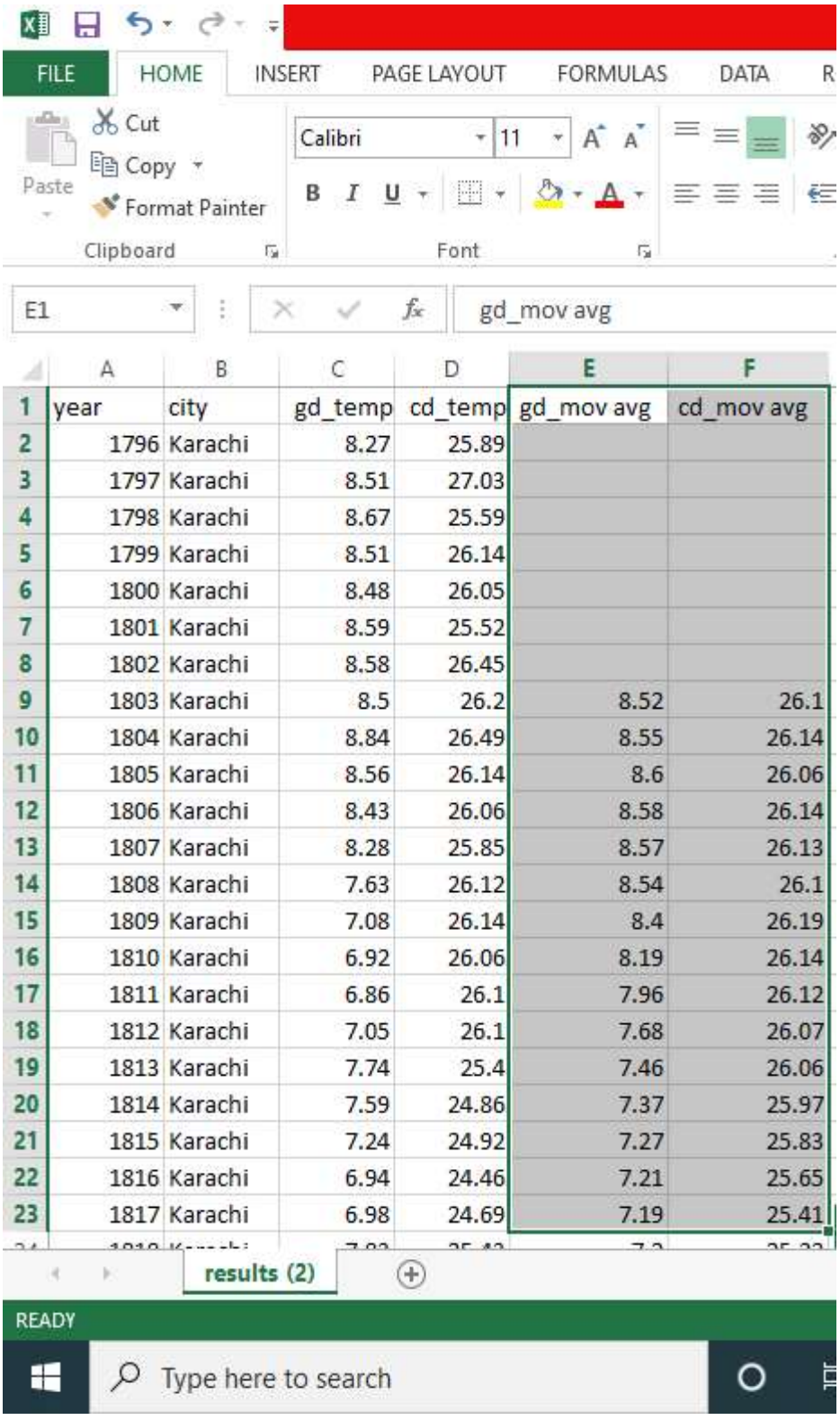
Because data was missing from different places in cd_temp column.



Calculating moving average (fourth step):

I calculated moving average of 7 years for my city and global temperature by using =AVERAGE (C2:C8) and =AVERAGE (D2:D8) formula respectively. Then I scrolled down to compute moving average for every cell.


results (2) - Excel (Product Activation Failed)
FILE
HOME
INSERT
PAGE LAYOUT
FORMULAS
DATA
REVIEW
VIEW
Cut
Copy
Format Painter
Paste
Clipboard
Font
Wrap Text
Merge & Center
Alignment
General
Number
Conditional Formatting
Format as Table
Cell Styles
Styles
Insert
Delete
Format
Cells
AutoSum
Fill
Clear
Sort & Filter
Find & Select
Editing
D2
=AVERAGE(D2:D8
year city gd_temp cd_temp gd_mov a cd_mov avg
1796 Karachi 8.27 25.89
1797 Karachi 8.51 27.03
1798 Karachi 8.67 25.59
1799 Karachi 8.51 26.14
1800 Karachi 8.48 26.05
1801 Karachi 8.59 25.52
1802 Karachi 8.58 26.45
1803 Karachi 8.5 26.2 =AVERAGE(D2:D8
AVERAGE(number1, [number2], ...)
1804 Karachi 8.84 26.49
1805 Karachi 8.56 26.14
1806 Karachi 8.43 26.06
1807 Karachi 8.28 25.85
1808 Karachi 7.63 26.12
1809 Karachi 7.08 26.14
1810 Karachi 6.92 26.06
1811 Karachi 6.86 26.1
1812 Karachi 7.05 26.1
1813 Karachi 7.74 25.4
1814 Karachi 7.59 24.86
1815 Karachi 7.24 24.92
1816 Karachi 6.94 24.46
1817 Karachi 6.98 24.69
Activate Windows
Go to Settings to activate Windows.
results (2)
POINT
100%
Type here to search
12:19 PM
9/2/2020

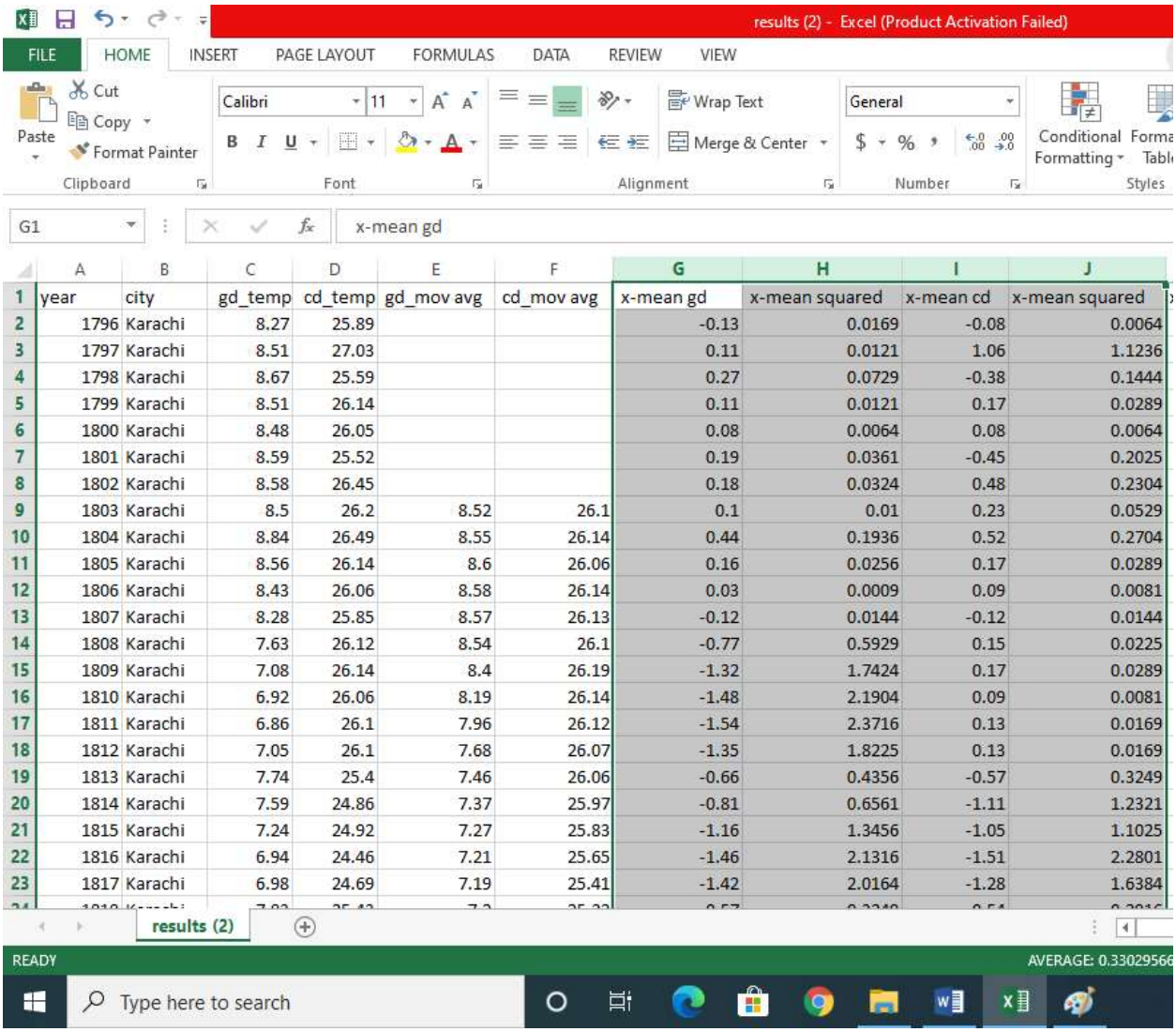


	A	B	C	D	E	F
1	year	city	gd_temp	cd_temp	gd_mov avg	cd_mov avg
2	1796	Karachi	8.27	25.89		
3	1797	Karachi	8.51	27.03		
4	1798	Karachi	8.67	25.59		
5	1799	Karachi	8.51	26.14		
6	1800	Karachi	8.48	26.05		
7	1801	Karachi	8.59	25.52		
8	1802	Karachi	8.58	26.45		
9	1803	Karachi	8.5	26.2	8.52	26.1
10	1804	Karachi	8.84	26.49	8.55	26.14
11	1805	Karachi	8.56	26.14	8.6	26.06
12	1806	Karachi	8.43	26.06	8.58	26.14
13	1807	Karachi	8.28	25.85	8.57	26.13
14	1808	Karachi	7.63	26.12	8.54	26.1
15	1809	Karachi	7.08	26.14	8.4	26.19
16	1810	Karachi	6.92	26.06	8.19	26.14
17	1811	Karachi	6.86	26.1	7.96	26.12
18	1812	Karachi	7.05	26.1	7.68	26.07
19	1813	Karachi	7.74	25.4	7.46	26.06
20	1814	Karachi	7.59	24.86	7.37	25.97
21	1815	Karachi	7.24	24.92	7.27	25.83
22	1816	Karachi	6.94	24.46	7.21	25.65
23	1817	Karachi	6.98	24.69	7.19	25.41

As can be seen in the screenshot above, moving average is calculated and values are inserted in column E and F respectively.

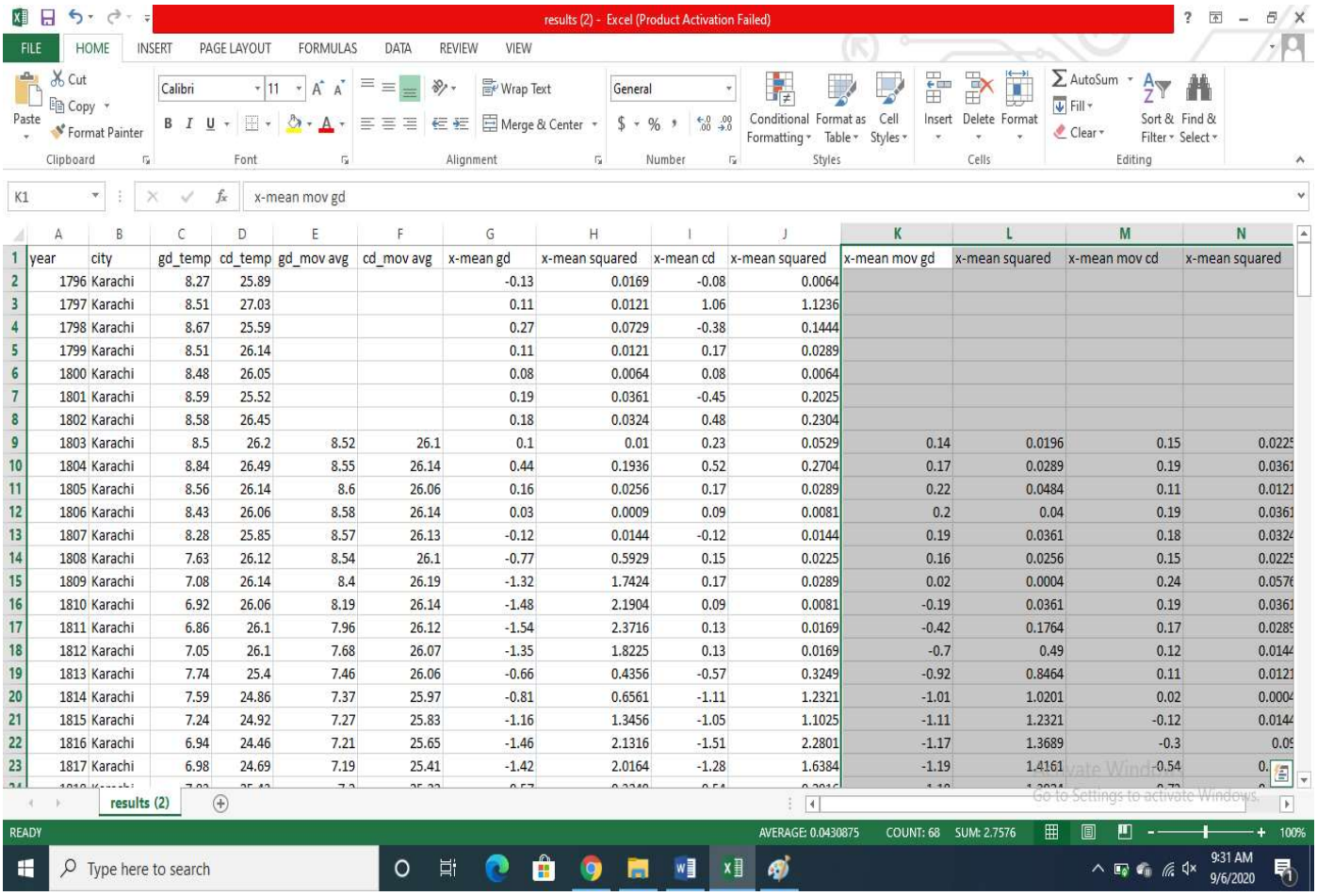
Mean, Variance and Standard deviation (optional step in this project):

First, I did calculations before taking moving average.



	A	B	C	D	E	F	G	H	I	J
1	year	city	gd_temp	cd_temp	gd_mov avg	cd_mov avg	x-mean gd	x-mean squared	x-mean cd	x-mean squared
2	1796	Karachi	8.27	25.89			-0.13	0.0169	-0.08	0.0064
3	1797	Karachi	8.51	27.03			0.11	0.0121	1.06	1.1236
4	1798	Karachi	8.67	25.59			0.27	0.0729	-0.38	0.1444
5	1799	Karachi	8.51	26.14			0.11	0.0121	0.17	0.0289
6	1800	Karachi	8.48	26.05			0.08	0.0064	0.08	0.0064
7	1801	Karachi	8.59	25.52			0.19	0.0361	-0.45	0.2025
8	1802	Karachi	8.58	26.45			0.18	0.0324	0.48	0.2304
9	1803	Karachi	8.5	26.2	8.52	26.1	0.1	0.01	0.23	0.0529
10	1804	Karachi	8.84	26.49	8.55	26.14	0.44	0.1936	0.52	0.2704
11	1805	Karachi	8.56	26.14	8.6	26.06	0.16	0.0256	0.17	0.0289
12	1806	Karachi	8.43	26.06	8.58	26.14	0.03	0.0009	0.09	0.0081
13	1807	Karachi	8.28	25.85	8.57	26.13	-0.12	0.0144	-0.12	0.0144
14	1808	Karachi	7.63	26.12	8.54	26.1	-0.77	0.5929	0.15	0.0225
15	1809	Karachi	7.08	26.14	8.4	26.19	-1.32	1.7424	0.17	0.0289
16	1810	Karachi	6.92	26.06	8.19	26.14	-1.48	2.1904	0.09	0.0081
17	1811	Karachi	6.86	26.1	7.96	26.12	-1.54	2.3716	0.13	0.0169
18	1812	Karachi	7.05	26.1	7.68	26.07	-1.35	1.8225	0.13	0.0169
19	1813	Karachi	7.74	25.4	7.46	26.06	-0.66	0.4356	-0.57	0.3249
20	1814	Karachi	7.59	24.86	7.37	25.97	-0.81	0.6561	-1.11	1.2321
21	1815	Karachi	7.24	24.92	7.27	25.83	-1.16	1.3456	-1.05	1.1025
22	1816	Karachi	6.94	24.46	7.21	25.65	-1.46	2.1316	-1.51	2.2801
23	1817	Karachi	6.98	24.69	7.19	25.41	-1.42	2.0164	-1.28	1.6384

Then, I did calculations after applying smoothing data filter i.e. moving average here.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	year	city	gd_temp	cd_temp	gd_mov avg	cd_mov avg	x-mean gd	x-mean squared	x-mean cd	x-mean squared	x-mean mov gd	x-mean squared	x-mean mov cd	x-mean squared
2	1796	Karachi	8.27	25.89			-0.13	0.0169	-0.08	0.0064				
3	1797	Karachi	8.51	27.03			0.11	0.0121	1.06	1.1236				
4	1798	Karachi	8.67	25.59			0.27	0.0729	-0.38	0.1444				
5	1799	Karachi	8.51	26.14			0.11	0.0121	0.17	0.0289				
6	1800	Karachi	8.48	26.05			0.08	0.0064	0.08	0.0064				
7	1801	Karachi	8.59	25.52			0.19	0.0361	-0.45	0.2025				
8	1802	Karachi	8.58	26.45			0.18	0.0324	0.48	0.2304				
9	1803	Karachi	8.5	26.2	8.52	26.1	0.1	0.01	0.23	0.0529	0.14	0.0196	0.15	0.0225
10	1804	Karachi	8.84	26.49	8.55	26.14	0.44	0.1936	0.52	0.2704	0.17	0.0289	0.19	0.0361
11	1805	Karachi	8.56	26.14	8.6	26.06	0.16	0.0256	0.17	0.0289	0.22	0.0484	0.11	0.0121
12	1806	Karachi	8.43	26.06	8.58	26.14	0.03	0.0009	0.09	0.0081	0.2	0.04	0.19	0.0361
13	1807	Karachi	8.28	25.85	8.57	26.13	-0.12	0.0144	-0.12	0.0144	0.19	0.0361	0.18	0.0324
14	1808	Karachi	7.63	26.12	8.54	26.1	-0.77	0.5929	0.15	0.0225	0.16	0.0256	0.15	0.0225
15	1809	Karachi	7.08	26.14	8.4	26.19	-1.32	1.7424	0.17	0.0289	0.02	0.0004	0.24	0.0576
16	1810	Karachi	6.92	26.06	8.19	26.14	-1.48	2.1904	0.09	0.0081	-0.19	0.0361	0.19	0.0361
17	1811	Karachi	6.86	26.1	7.96	26.12	-1.54	2.3716	0.13	0.0169	-0.42	0.1764	0.17	0.0289
18	1812	Karachi	7.05	26.1	7.68	26.07	-1.35	1.8225	0.13	0.0169	-0.7	0.49	0.12	0.0144
19	1813	Karachi	7.74	25.4	7.46	26.06	-0.66	0.4356	-0.57	0.3249	-0.92	0.8464	0.11	0.0121
20	1814	Karachi	7.59	24.86	7.37	25.97	-0.81	0.6561	-1.11	1.2321	-1.01	1.0201	0.02	0.0004
21	1815	Karachi	7.24	24.92	7.27	25.83	-1.16	1.3456	-1.05	1.1025	-1.11	1.2321	-0.12	0.0144
22	1816	Karachi	6.94	24.46	7.21	25.65	-1.46	2.1316	-1.51	2.2801	-1.17	1.3689	-0.3	0.09
23	1817	Karachi	6.98	24.69	7.19	25.41	-1.42	2.0164	-1.28	1.6384	-1.19	1.4161	-0.54	0.2

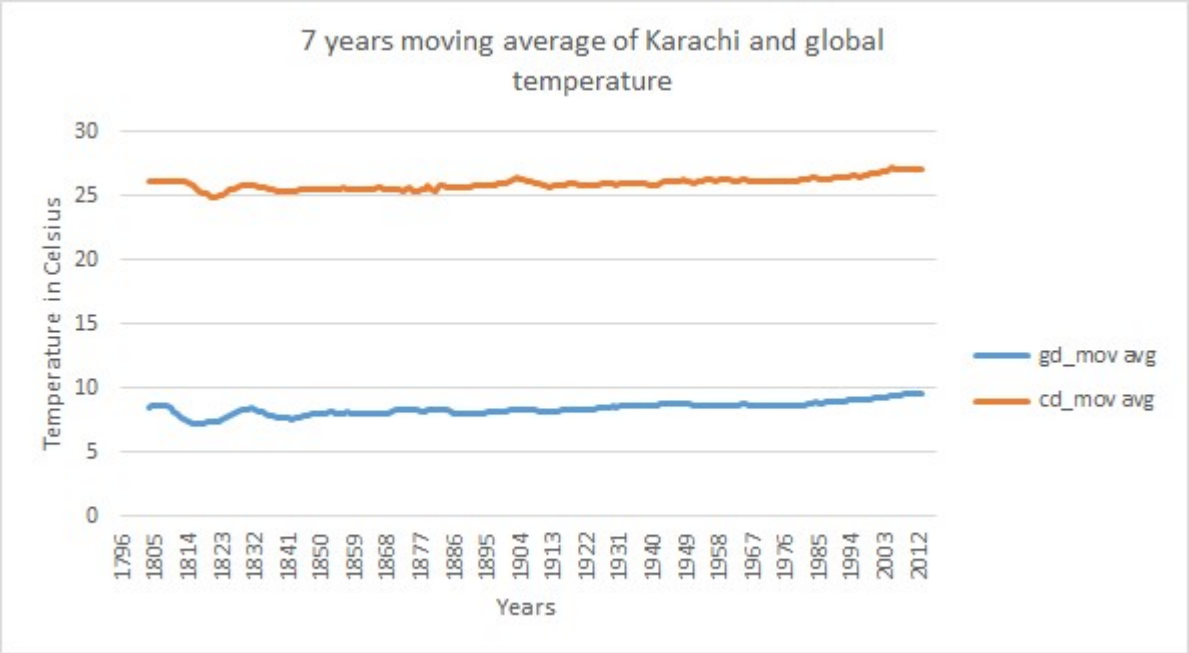
Results:

Second and third columns have global and city data calculation results (raw data) and fourth and fifth columns contain moving average calculations.

Mean:	8.4	25.97	8.38	25.95
Variance:	0.30	0.36	0.23	0.19
Standard Deviation:	0.55	0.60	0.48	0.44

results (2)

Line chart of 7 year moving average for city and global data (fifth step):



Observations:

According to line chart, following observations may be deduced:

1. Karachi is 18 degrees hotter than global temperature.

2. Karachi's temperature is more volatile as compared to global.

3. Global temperature is gradually increasing.

4. There is significant rise in global temperature between 1990 and 2013. Temperature has been in the range of 7 and 8 degree Celsius before. In 1990, it was recorded 8.9 and has been increasing since. For the year 2013, it was recorded 9.56 degree Celsius.

5. Noticeable change in Karachi's temperature is between 2000 and 2013 as temperature was increased from 26 degrees to 27 degrees Celsius. Before 2000, temperature was in range of 25 and 26 degree Celsius.

6. Correlation coefficient between Karachi and global temperature is 0.71 which suggests strong positive correlation between two variables.

7. The average global temperature is 8.38 while the average temperature for Karachi is 25.95

8. Temperature difference (delta t) of global temperature is 1.34 and Karachi is 1.32