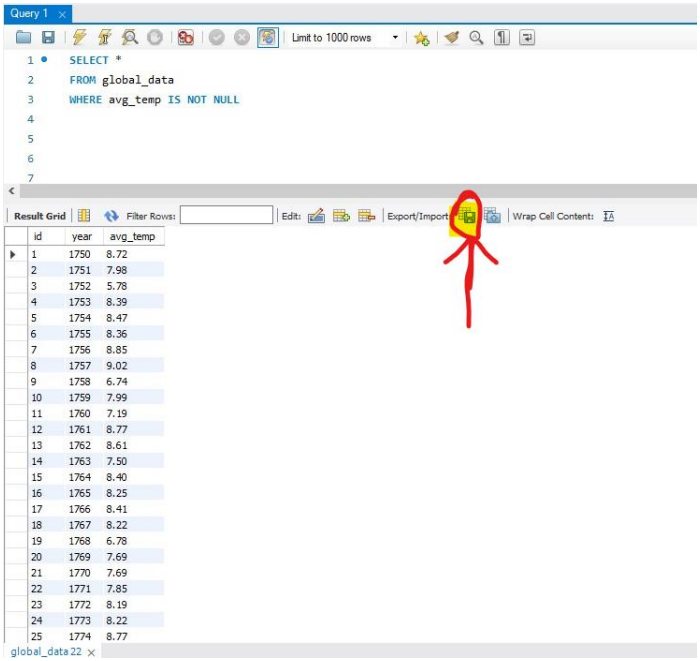
Tony Jin Python for Data Science HW 1

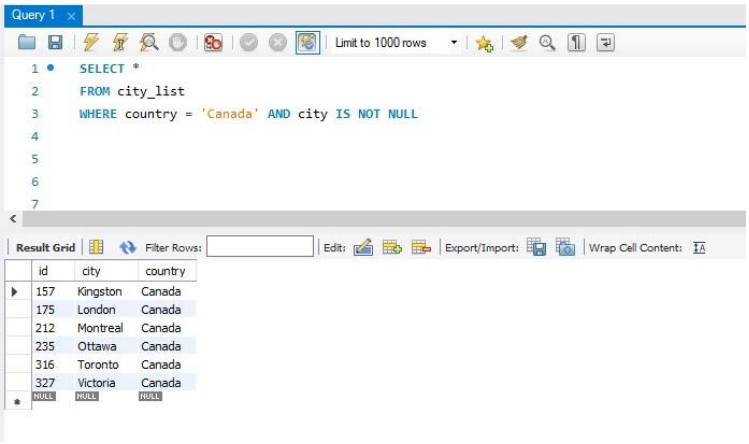
Step 1: Access Global Average Data and Export to CSV for import into Excel (make sure there are no NULL values)



```
SELECT *
FROM global_data
WHERE avg_temp IS NOT NULL
```

id	year	avg_temp
1	1750	8.72
2	1751	7.98
3	1752	5.78
4	1753	8.39
5	1754	8.47
6	1755	8.36
7	1756	8.85
8	1757	9.02
9	1758	6.74
10	1759	7.99
11	1760	7.19
12	1761	8.77
13	1762	8.61
14	1763	7.50
15	1764	8.40
16	1765	8.25
17	1766	8.41
18	1767	8.22
19	1768	6.78
20	1769	7.69
21	1770	7.69
22	1771	7.85
23	1772	8.19
24	1773	8.22
25	1774	8.77

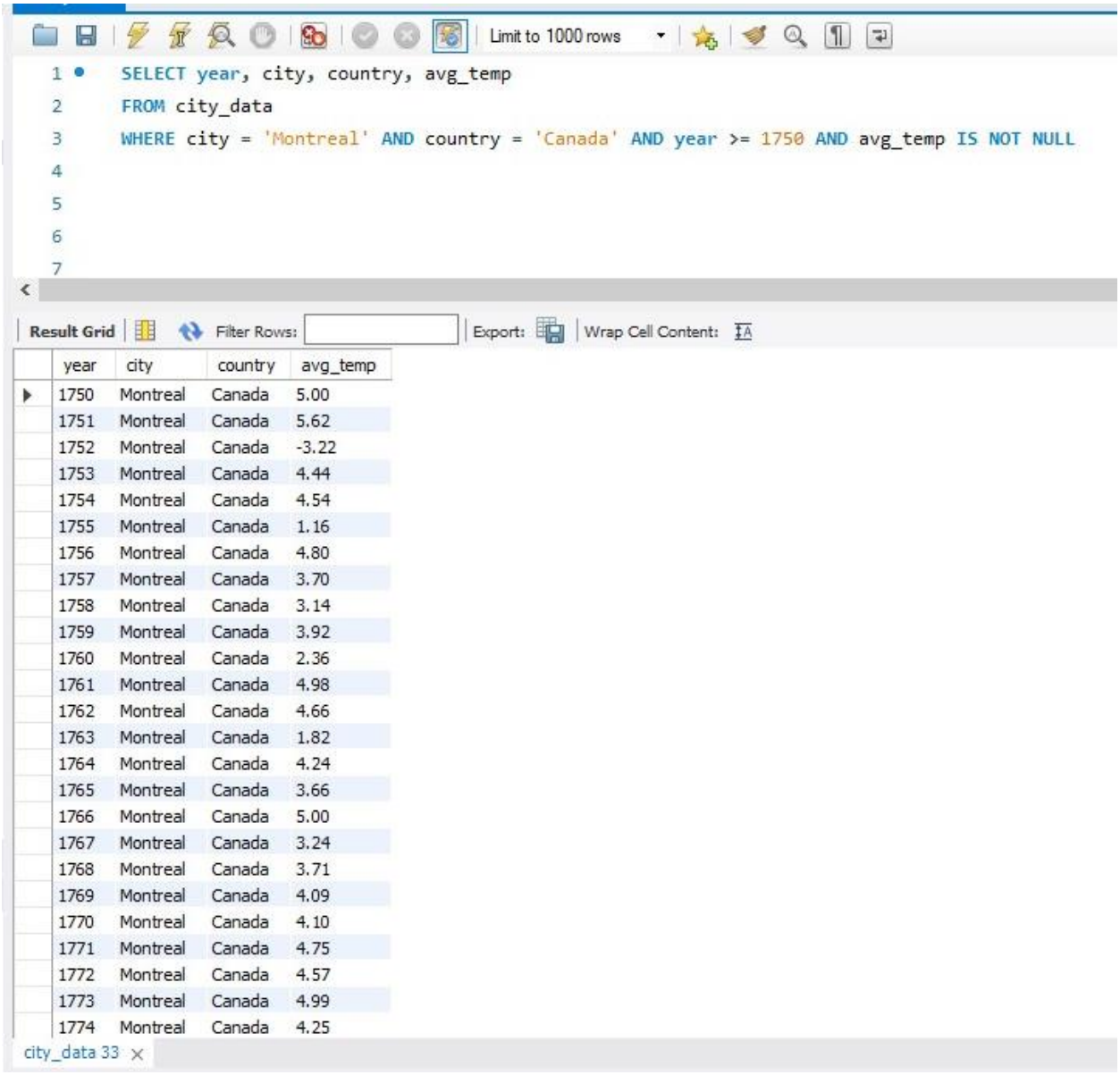
Step 2: consult city_list to determine which Canadian cities are on the list for me to use



```
SELECT *
FROM city_list
WHERE country = 'Canada' AND city IS NOT NULL
```

id	city	country
157	Kingston	Canada
175	London	Canada
212	Montreal	Canada
235	Ottawa	Canada
316	Toronto	Canada
327	Victoria	Canada

Step 3: Select for City Data, removing null values. Export to CSV. I will be doing this for Montreal because it is a neat city! I noticed that Montreal's earliest timepoint is 1743 while global average starts in 1750. Steps were taken to eliminate data points with no valid comparison.



```
SELECT year, city, country, avg_temp
FROM city_data
WHERE city = 'Montreal' AND country = 'Canada' AND year >= 1750 AND avg_temp IS NOT NULL
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

year	city	country	avg_temp
1750	Montreal	Canada	5.00
1751	Montreal	Canada	5.62
1752	Montreal	Canada	-3.22
1753	Montreal	Canada	4.44
1754	Montreal	Canada	4.54
1755	Montreal	Canada	1.16
1756	Montreal	Canada	4.80
1757	Montreal	Canada	3.70
1758	Montreal	Canada	3.14
1759	Montreal	Canada	3.92
1760	Montreal	Canada	2.36
1761	Montreal	Canada	4.98
1762	Montreal	Canada	4.66
1763	Montreal	Canada	1.82
1764	Montreal	Canada	4.24
1765	Montreal	Canada	3.66
1766	Montreal	Canada	5.00
1767	Montreal	Canada	3.24
1768	Montreal	Canada	3.71
1769	Montreal	Canada	4.09
1770	Montreal	Canada	4.10
1771	Montreal	Canada	4.75
1772	Montreal	Canada	4.57
1773	Montreal	Canada	4.99
1774	Montreal	Canada	4.25

city_data 33

Step 4: After CSV data was loaded into EXCEL, I copied relevant columns to a new sheet and calculated 5 year and 10 year moving averages for average temperature.

	A	B	C	D	E	F	G	H	I	J
1	id	year	avg_temp	5 Year MA	10 Year MA	city	avg_temp(MTL)	5 Year MA	10 Year MA	
2	1	1750	8.72			Montreal	5.00			
3	2	1751	7.98			Montreal	5.62			
4	3	1752	5.78			Montreal	-3.22			
5	4	1753	8.39			Montreal	4.44			
6	5	1754	8.47	7.87		Montreal	4.54	3.28		
7	6	1755	8.36	7.80		Montreal	1.16	2.51		
8	7	1756	8.85	7.97		Montreal	4.80	2.34		
9	8	1757	9.02	8.62		Montreal	3.70	3.73		
10	9	1758	6.74	8.29		Montreal	3.14	3.47		
11	10	1759	7.99	8.19	8.03	Montreal	3.92	3.34	=AVERAGE(G2:G11	
12	11	1760	7.19	7.96	7.877	Montreal	2.36	3.58		
13	12	1761	8.77	7.94	7.956	Montreal	4.98	3.62		
14	13	1762	8.61	7.86	8.239	Montreal	4.66	3.81		
15	14	1763	7.5	8.01	8.15	Montreal	1.82	3.55		
16	15	1764	8.4	8.09	8.143	Montreal	4.24	3.61		
17	16	1765	8.25	8.31	8.132	Montreal	3.66	3.87		
18	17	1766	8.41	8.23	8.088	Montreal	5.00	3.88		
19	18	1767	8.22	8.16	8.008	Montreal	3.24	3.59		
20	19	1768	6.78	8.01	8.012	Montreal	3.71	3.97		
21	20	1769	7.69	7.87	7.982	Montreal	4.09	3.94		
22	21	1770	7.69	7.76	8.032	Montreal	4.10	4.03		
23	22	1771	7.85	7.65	7.94	Montreal	4.75	3.98		
24	23	1772	8.19	7.64	7.898	Montreal	4.57	4.24		
25	24	1773	8.22	7.93	7.97	Montreal	4.99	4.50		
26	25	1774	8.77	8.14	8.007	Montreal	4.25	4.53		
27	26	1775	9.18	8.44	8.1	Montreal	5.26	4.76		
28	27	1776	8.3	8.53	8.089	Montreal	4.32	4.68		
29	28	1777	8.26	8.55	8.093	Montreal	3.81	4.53		
30	29	1778	8.54	8.61	8.269	Montreal	1.71	3.87		
31	30	1779	8.98	8.65	8.398	Montreal	4.87	3.99		
32	31	1780	9.43	8.70	8.572	Montreal	7.47	4.44		
33	32	1781	8.1	8.66	8.597	Montreal	4.66	4.50		
34	33	1782	7.9	8.59	8.568	Montreal	4.07	4.56		
35	34	1783	7.68	8.42	8.514	Montreal	3.77	4.97		
36	35	1784	7.86	8.19	8.423	Montreal	3.49	4.69		

Montreal Temp AVG | Global Temp AVG | **MTL_Global Comparison**

Step 5: Excel's Native chart creator is adequate for making line graphs

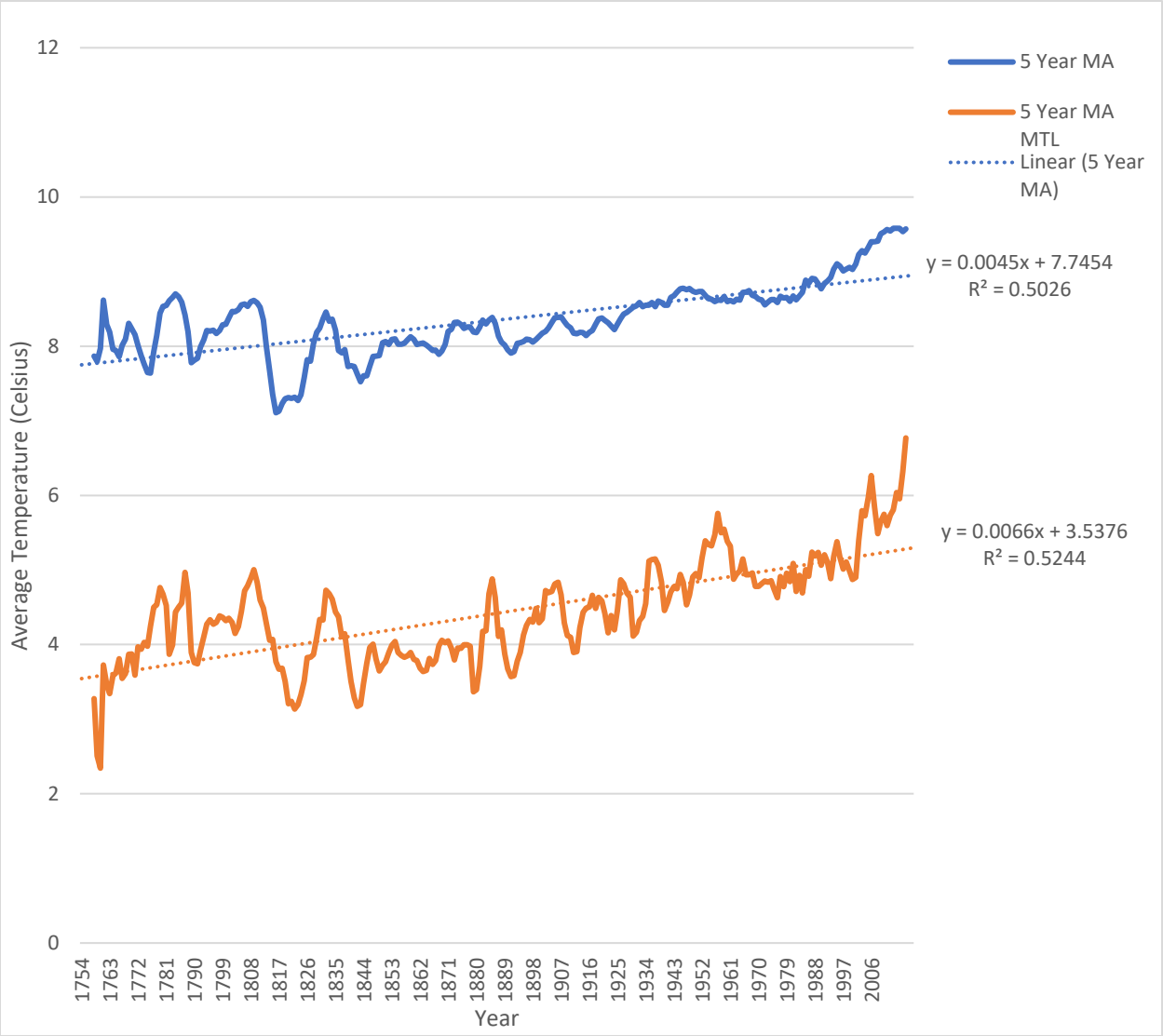


Figure 1: Comparison of Average temperature between Montreal (Orange) and Global data (Blue) in 5 year moving averages. Trendlines were fitted for Global data and Montreal Data with R^2 values of 0.5026 and 0.5244 respectively.

Observations:

- The average temperature trends for Montreal is more or less moving with the global trend. **Both have been increasing over this period.** (Show this to people who don't think Global Warming is happening!)
- Fun fact: The sharp decrease in 1816 is known as the year without a summer. This is likely due to a series of massive volcanic eruptions around the world.
- Montreal data seems to experience much more pronounced peaks and troughs in temperature data than overall global data

- The slope for the linear trendline for the MTL data is steeper, indicating a greater rate of increase in average temperature over time
- Interestingly, the MTL trendline experiences a greater R^2 value, indicating a slightly better fit. Maybe due to the fact that all temperature is from one location? Maybe the difference is negligible? Shouldn't it be lower due to the larger variations? I don't have a stats background so I don't really know.

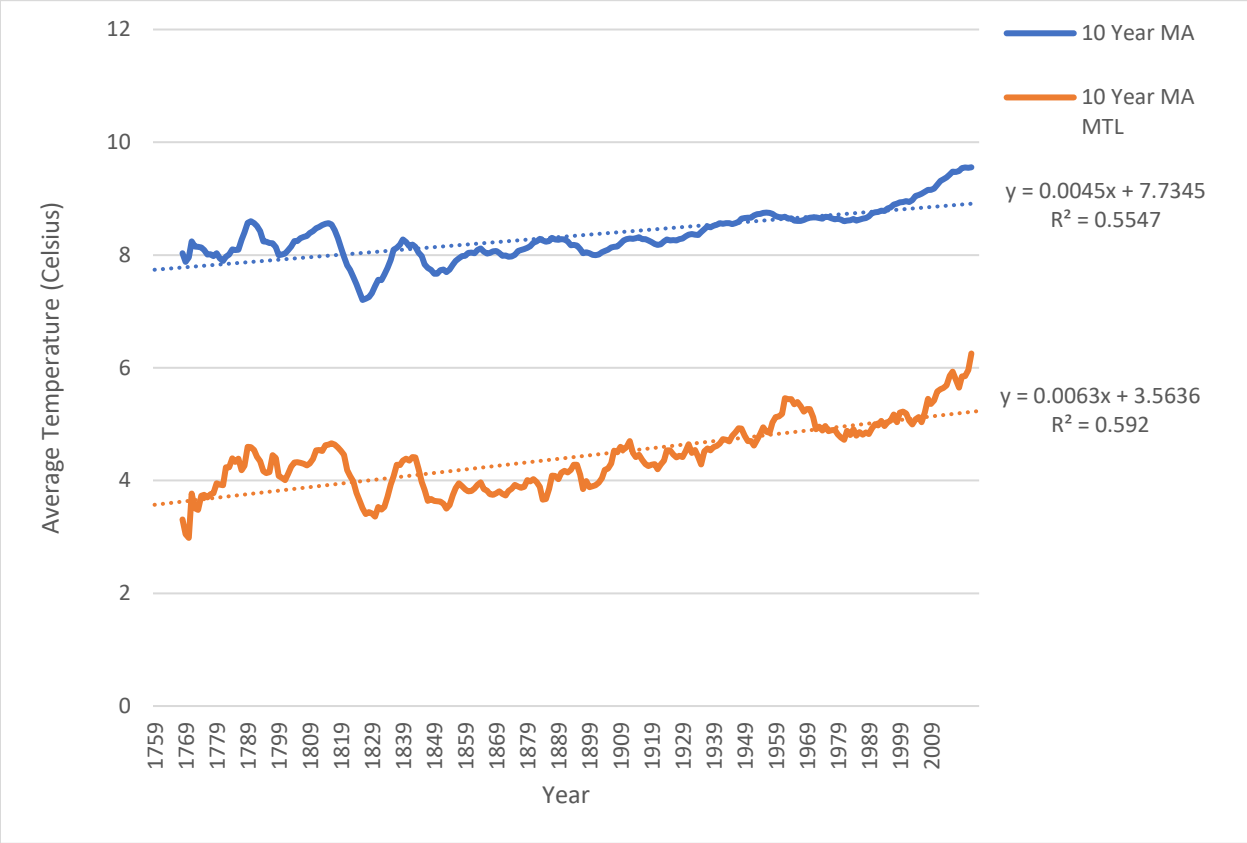


Figure 2: Comparison of Average temperature between Montreal (Orange) and Global data (Blue) in 10 year moving averages. Trendlines were fitted for Global data and Montreal Data with R^2 values of 0.5547and 0.5920 respectively.

Observations

- The same general trends can be observed
- As expected, the extreme peaks and troughs from Montreal in the 5 year MA have been smoothed over somewhat.
- The R^2 Values are higher now with the smoother graphs, but the higher R^2 value is still seen in the Montreal Data.