Exploring Weather Trends Project

By

Student:	Devin Datt
Email:	ddatt@rogers.com
Date:	December 18th, 2017
Submission:	version 1.0

Summary

To analyze local and global temperature data and compare the temperature trends from a local city (ie. Toronto) to overall global temperature trends and provide insights on differences and similarities.

Method

A) Extract data: Access the Udacity data page to the temperature data sources (see Appendix A) and export both temperature data for global stats and local City of Toronto stats. Used the following SQL commands to interact and extract the relative data that took the form below:

Local Temperature Extraction:

STEP 1: Ran following SQL command to locate if city of "Toronto, Canada" is in dataset table of 'city_list':

```
SELECT *
FROM city_list
WHERE city = 'Toronto';

RESULT:  1 record found (city = Toronto, country = Canada)
```

STEP 2: Once found Toronto is in the list, ran following SQL command to then extract all data (the year and its average yearly temperature) for that city in the 'city_data' table:

```
SELECT *
FROM city_data
WHERE city = 'Toronto';

RESULT:  271 records found (downloaded to CSV file)
```

Global Temperature Extraction:

STEP 3: Once I extracted the local city data, next task was to filter for all global temperature data. I ran the following SQL command to query the 'global_data' table:

```
SELECT *
FROM global_data;

RESULT:  266 records found (downloaded to CSV file)
```

STEP 4: As mentioned above, after pulling the pertinent data needed, I exported the datasets into two separate CSV files (named 'Toronto_Temp_Data.csv' and 'Global_Temp_Data.csv') to my local computer. See Appendix B & C respectively for summarized raw data)

B) Massage data:

STEP 1: Pull Data Together: Using MS Excel I copied each CSV raw data into its own separate tabs into 3rd sheet (called 'Weather_Trend_data_v1.xlsx'). Then combined all the data into a 3rd tab (named 'City_Global Data').

STEP 2: Aligning Data: Since city data contained more year data vs. the global (ie. data for years 1743-1755 existed in Toronto data then global data) I aligned the dates to match and removed the irrelevant columns (city, country, year…etc) only leaving 3 columns (year, City_avg_temp, and Global_avg_temp). I created two additional columns for each data set to calculate their respective 7 & 14 Day Moving Averages bringing the total columns in spreadsheet to seven (see Figure 1.0). Decision was made to only plot the data from years 1763 to 2013 due to the fact the first full data point for the 14-MA stat for both city and global sources needed to be calculated from years 1750-1763 as illustrated below:

year	City_avg_temp	City_temp_7MA	City_temp_14MA	Global_avg_temp	Global_temp_7MA	Global_temp_14MA
1743	-0.11					
1744	8.38					
1745	-3.96					
1746						
1747						
1748						
1749		1.44				
1750	6.29	3.57		8.72	8.72	
1751	6.84	3.06		7.98	8.35	
1752	-1.1	4.01		5.78	7.49	
1753	5.76	4.45		8.39	7.72	
1754	5.94	4.75		8.47	7.87	
1755	2.81	4.42		8.36	7.95	
1756	6.37	4.70	3.72	8.85	8.08	8.08
1757	5.13	4.54	4.25	9.02	8.12	8.20
1758	4.37	4.18	3.85	6.74	7.74	8.03
1759	5.27	5.[illegible]	4.77	7.99	8.26	8.03
1760	3.74	4.80	4.67	7.19	8.09	7.95
1761	6.25	4.85	4.81	8.77	8.13	8.02
1762	5.79	5.27	4.88	8.61	8.17	8.07
1763	3.32	4.84	4.77	7.5	7.97	8.03
1764	5.76	4.93	4.73	8.4	7.89	8.00
1765	5.14	5.04	4.61	8.25	8.10	8.02
1766	6.27	5.18	5.14	8.41	8.16	8.21
1767	4.74	5.32	5.06	8.22	8.31	8.20
1768	4.81	5.12	4.98	6.78	8.02	8.08
1769	5.44	5.07	5.17	7.69	7.89	8.03
1770	5.35	5.36	5.10	7.69	7.92	7.95

Figure 1.0: Portion of the Excel data of the City of Toronto, Global yearly average temperatures showing why the 7-MA (Blue box) and the 14-MA (Red box) was best to use 1763 as the start of year to plot data. The was also data that was omitted at the beginning to better align data for comparison (Green Box).

STEP 3: Make Data Readable: Created additional columns to calculate 7 and 14 Day Moving average, using the Excel AVERAGE formula to find the average temperature from the previous 7 and 14 years (ie. =AVERAGE (B2:B8) & AVERAGE (B2:B15) respectively) This was considered in order to smooth out the data to better see trends and patterns when comparing the Local and Global temperatures in a chart (which is explained in the next step).

C) Charting Data:

STEP 1: Chart Raw Data: To analyze the data in the CSV files I used MS Excel to read and analyze the info. After downloading the data properly in their respective excel sheets I combined the two datasets into one chart for better management and analysis. This can be seen as the chart in Figure 2.0 below:

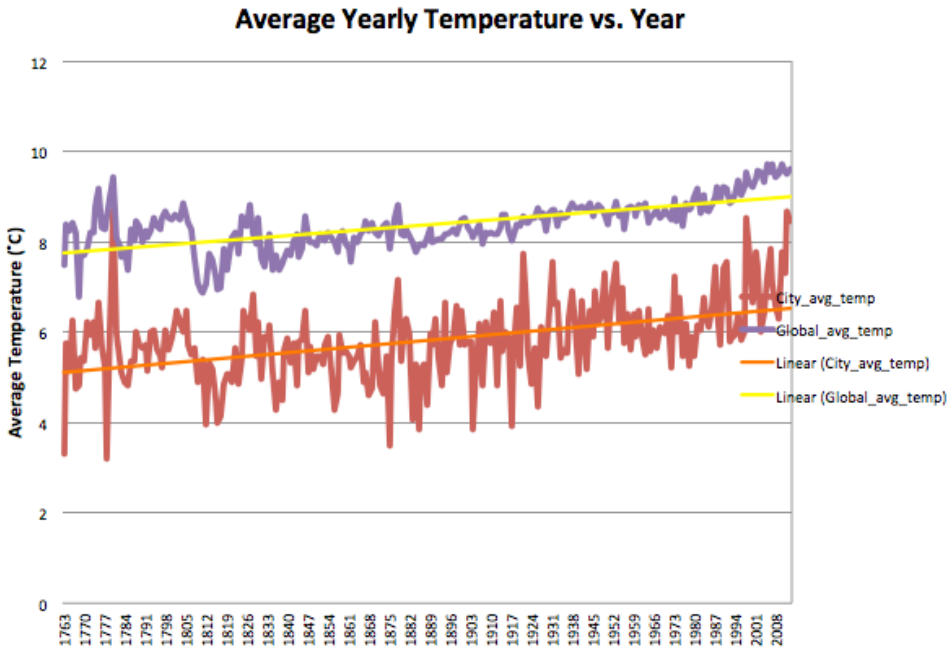


Figure 2.0: Chart comparing the raw data of average temperatures of both the city of Toronto (Red) and its respective Global yearly data with their respective trendlines moving up.

STEP 2: Chart Moving Averages: I then charted the Moving Average (MA) for both 7 and 14 days using the separate columns as mentioned in the previous Section B: Step 3 (Massage Data: Making Data Readable). This produced the following charts in Figure 3.0 (7-Day MA) and Figure 4.0 (14-Day MA) below:

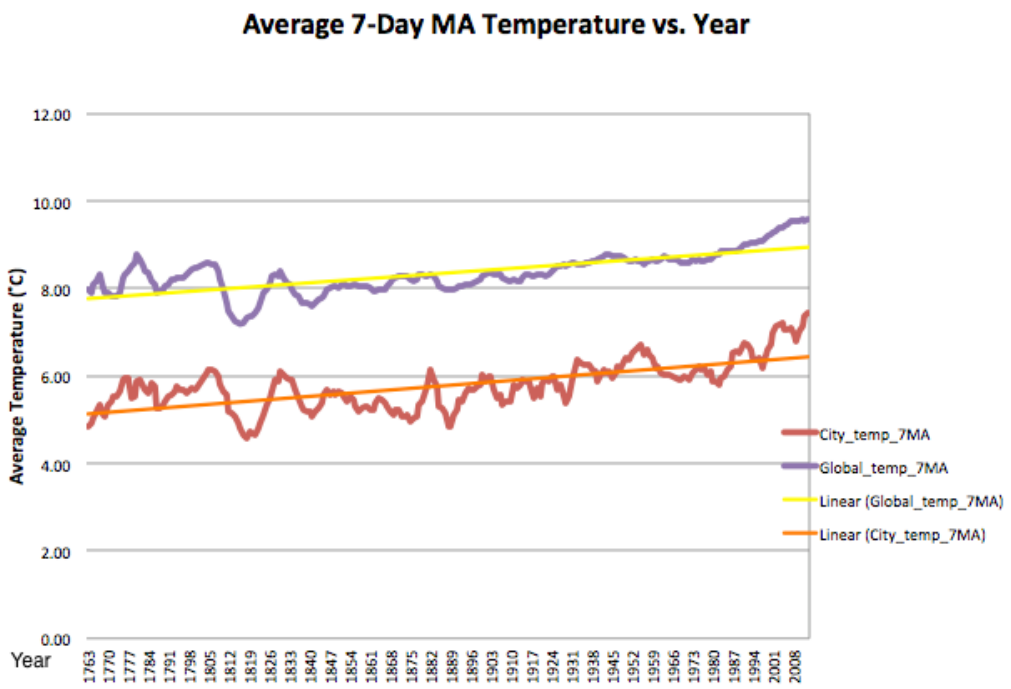


Figure 3.0: Chart comparing the 7-Day Moving Average (MA) temperatures of both the city of Toronto (Red) and its respective Global yearly data with their respective trendlines.

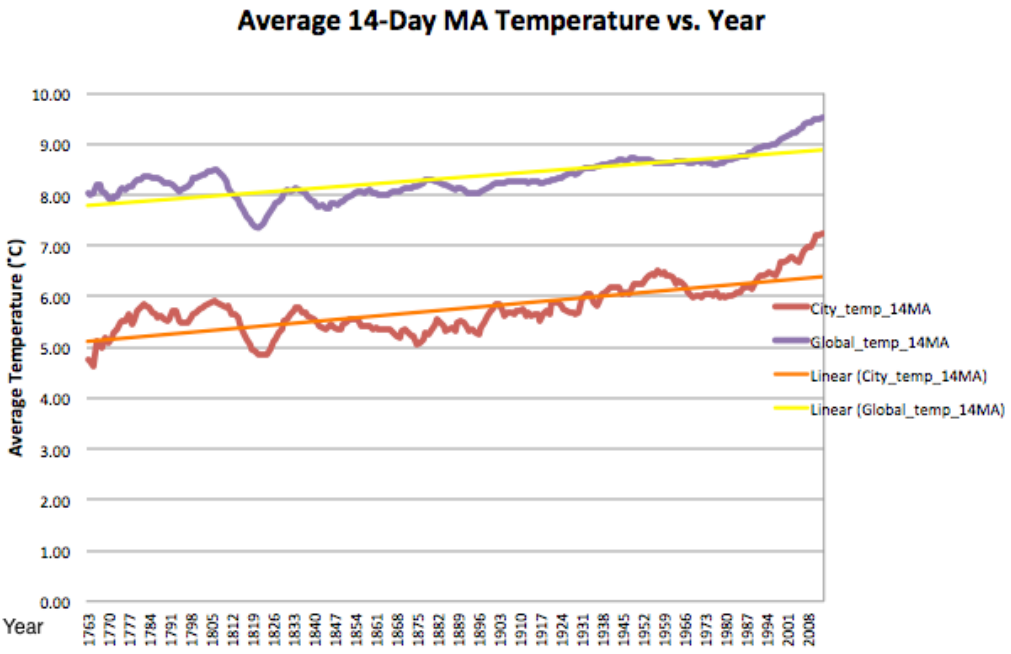


Figure 4.0: Chart comparing the 14-Day Moving Average (MA) temperatures of both the city of Toronto (Red) and its respective Global yearly data.

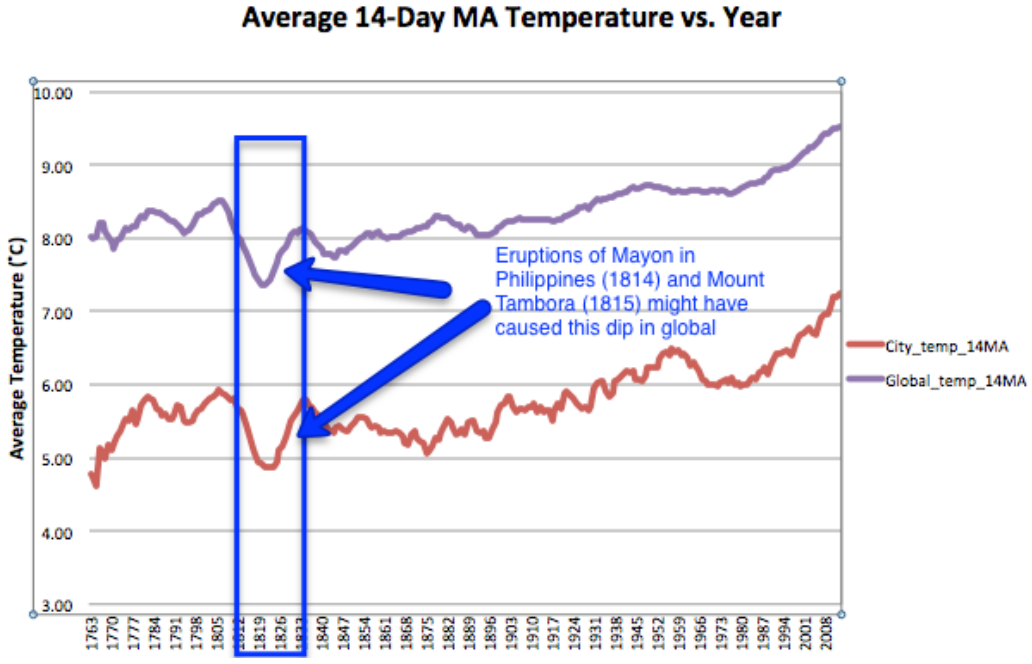
Observations

Based on the charts of the two data plots we can see charts smooth out as we move from Figure 2.0 to Figure 4.0. Due to this fact we will use Figure 4.0 (using 14-MA) to best draw our observations below. Key observations noticed are:

1) Most (of not all) of all data points are above 4˚C for the city temperatures and above 6˚C for the global temperatures. This can be do to the fact that Toronto is in the northern latitudes and experiences more average colder weather.

2) Most of the weather seems to tract around the same temperature for both data sets from beginning up to near late 1800's (in Toronto ~ 5.5˚C and Global ~8˚C). After which point we notice a definite increase form the 1870's to present day. This is most likely explained to the increase use of fossil fuels in the beginning of the Industrial Revolution which were increasingly pumped into the atmosphere causing more of a "greenhouse" affect and affecting all cities globally.

3) There is a very distinct dip in city and global average temperatures around 1814 to 1820's. This might be due to a 'volcanic winter' caused by several volcanic events (see Appendix A: "Year Without a Summer – Wikipedia)



4) Since mid 1700's the Global average temperature has been consistently above Toronto's yearly average.

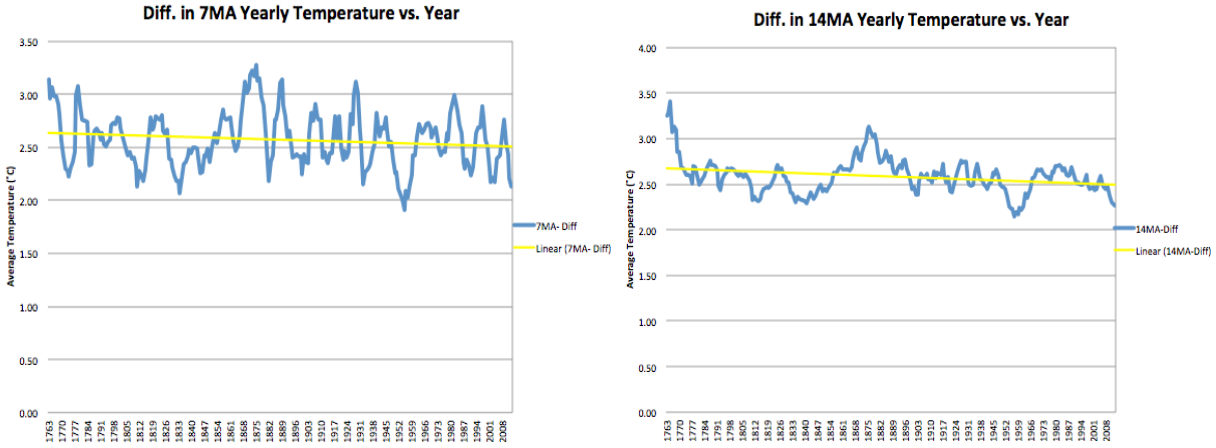
5) Global temperature has less variation then the local city especially after the 1850's. Possible causes of this are:

- The increase is aggregated data from other cities

- More precise measuring systems used

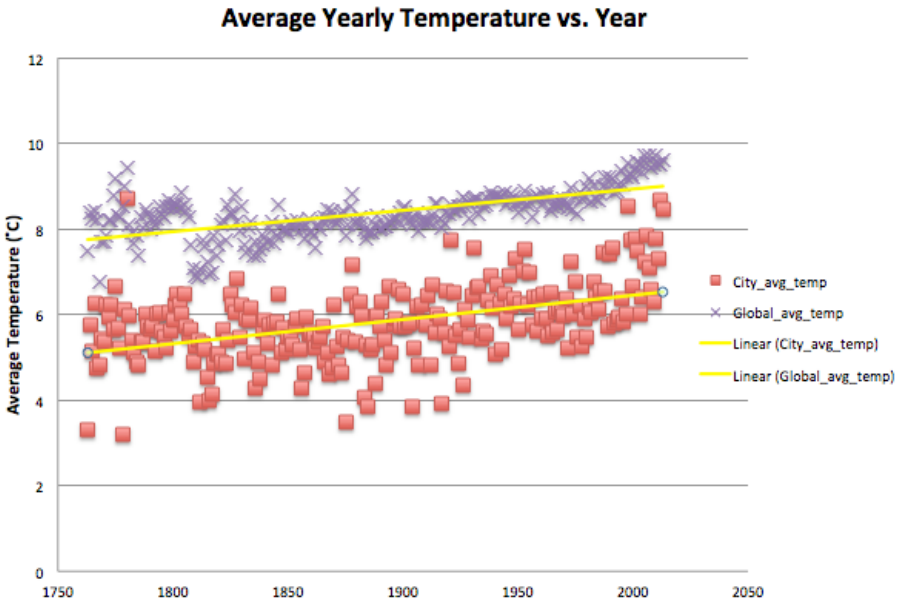
6) The City of Toronto's average yearly temperature varies greatly between 4-7 ˚C

7) If we look at the difference in the 7-MA and 14-MA data between both temperatures we notice a trendline that is decreasing slightly:



This decrease means the city average temperature is converging with global average temperatures. A possible reason for this might be Toronto winters is decreasing and its summers are increasing.

8) We can describe the correlation coefficient (CC) of a scatter plot of average temperatures for both datasets as the Global CC been positive because and we described before in Figure 2.0 we have trendlines moving up and should be between 0.5 & 0.75. And the city of Toronto's temperature CC been still positive but less then the Globals (around 0.4-0.6) as a trend doesn't pass thru as many plot points:



Appendix A

Temperature Data Sources - Udacity Weather Trends access data page: https://classroom.udacity.com/nanodegrees/nd002/parts/ca2cdcb3-c3df-428a-92e7-8b2630c7549d/modules/188c878c-5365-4bf3-9fa8-08cf57336fc4/lessons/dce89631-d141-4a36-b3fd-5e8ec038bc70/concepts/530f21c0-2f37-4390-aaab-3ce440e56d80

"Year Without a Summer" - https://en.wikipedia.org/wiki/Year_Without_a_Summer

Appendix B

Raw Local City Yearly Average Temperature Data (for City = Toronto):

year	avg_temp	1802	6.46	1842	5.75
1763	3.32	1803	6.23	1843	4.81
1764	5.76	1804	6.02	1844	5.8
1765	5.14	1805	6.49	1845	5.81
1766	6.27	1806	5.7	1846	6.47
1767	4.74	1807	5.49	1847	5.11
1768	4.81	1808	5.66	1848	5.68
1769	5.44	1809	4.9	1849	5.18
1770	5.35	1810	5.24	1850	5.45
1771	6.23	1811	5.38	1851	5.45
1772	5.93	1812	3.97	1852	5.27
1773	6.22	1813	5.31	1853	5.7
1774	5.64	1814	5.16	1854	5.9
1775	6.64	1815	4.52	1855	5.16
1776	5.69	1816	4	1856	4.26
1777	5.23	1817	4.12	1857	4.64
1778	3.18	1818	4.84	1858	5.93
1779	6.1	1819	5.07	1859	5.54
1780	8.7	1820	5.08	1860	5.58
1781	5.96	1821	4.88	1861	5.45
1782	5.39	1822	5.64	1862	5.21
1783	5.11	1823	4.86	1863	5.36
1784	4.9	1824	5.38	1864	5.47
1785	4.83	1825	6.46	1865	5.7
1786	5.35	1826	6.23	1866	4.88
1787	5.37	1827	6.03	1867	5.11
1788	6	1828	6.82	1868	4.59
1789	5.7	1829	5.46	1869	4.74
1790	5.64	1830	6.23	1870	6.22
1791	5.72	1831	4.96	1871	5.26
1792	5.15	1832	5.9	1872	4.83
1793	6.02	1833	5.88	1873	4.64
1794	6.05	1834	6.16	1874	5.47
1795	5.6	1835	5.11	1875	3.47
1796	5.48	1836	4.27	1876	5.41
1797	5.21	1837	4.89	1877	6.48
1798	6.03	1838	4.48	1878	7.15
1799	5.62	1839	5.62	1879	5.36
1800	5.85	1840	5.88	1880	6.19
1801	6.23	1841	5.33	1881	6.31

1882	5.97	1930	6.44	1978	5.24
1883	4.06	1931	7.56	1979	5.9
1884	5.29	1932	6.62	1980	5.48
1885	3.83	1933	6.66	1981	6.16
1886	5.17	1934	5.44	1982	6.04
1887	5.3	1935	5.6	1983	6.78
1888	4.39	1936	5.55	1984	6.42
1889	5.96	1937	6.32	1985	6.13
1890	5.71	1938	6.91	1986	6.56
1891	6.28	1939	6.14	1987	7.46
1892	5.4	1940	5.06	1988	6.56
1893	4.83	1941	6.68	1989	5.7
1894	6.64	1942	6.26	1990	7.41
1895	5.12	1943	5.19	1991	7.55
1896	5.75	1944	6.49	1992	5.79
1897	5.89	1945	5.9	1993	5.87
1898	6.59	1946	6.9	1994	5.95
1899	5.7	1947	6.18	1995	6.38
1900	6.47	1948	6.38	1996	5.81
1901	5.7	1949	7.31	1997	6
1902	5.8	1950	5.64	1998	8.54
1903	5.79	1951	6.26	1999	7.75
1904	3.85	1952	7.01	2000	6.67
1905	5.21	1953	7.51	2001	7.76
1906	6.19	1954	6.24	2002	7.48
1907	4.83	1955	6.97	2003	6.02
1908	6.22	1956	5.75	2004	6.4
1909	5.83	1957	6.42	2005	7.22
1910	5.77	1958	5.62	2006	7.85
1911	6.43	1959	6.4	2007	7.07
1912	4.82	1960	5.9	2008	6.58
1913	6.68	1961	6.46	2009	6.28
1914	5.59	1962	5.89	2010	7.77
1915	6	1963	5.49	2011	7.3
1916	5.89	1964	6.5	2012	8.66
1917	3.91	1965	5.59	2013	8.46
1918	5.57	1966	6.03		
1919	6.54	1967	5.66		
1920	5.24	1968	6.11		
1921	7.75	1969	6.03		
1922	6.5	1970	5.98		
1923	5.49	1971	6.37		
1924	4.86	1972	5.22		
1925	5.63	1973	7.23		
1926	4.34	1974	6.01		
1927	6.1	1975	6.75		
1928	5.86	1976	5.47		
1929	5.47	1977	6.19		

Appendix C

Raw Global Yearly Average Temperature Data:

year	avg_temp	1791	8.23	1833	8.01
1750	8.72	1792	8.09	1834	8.15
1751	7.98	1793	8.23	1835	7.39
1752	5.78	1794	8.53	1836	7.7
1753	8.39	1795	8.35	1837	7.38
1754	8.47	1796	8.27	1838	7.51
1755	8.36	1797	8.51	1839	7.63
1756	8.85	1798	8.67	1840	7.8
1757	9.02	1799	8.51	1841	7.69
1758	6.74	1800	8.48	1842	8.02
1759	7.99	1801	8.59	1843	8.17
1760	7.19	1802	8.58	1844	7.65
1761	8.77	1803	8.5	1845	7.85
1762	8.61	1804	8.84	1846	8.55
1763	7.5	1805	8.56	1847	8.09
1764	8.4	1806	8.43	1848	7.98
1765	8.25	1807	8.28	1849	7.98
1766	8.41	1808	7.63	1850	7.9
1767	8.22	1809	7.08	1851	8.18
1768	6.78	1810	6.92	1852	8.1
1769	7.69	1811	6.86	1853	8.04
1770	7.69	1812	7.05	1854	8.21
1771	7.85	1813	7.74	1855	8.11
1772	8.19	1814	7.59	1856	8
1773	8.22	1815	7.24	1857	7.76
1774	8.77	1816	6.94	1858	8.1
1775	9.18	1817	6.98	1859	8.25
1776	8.3	1818	7.83	1860	7.96
1777	8.26	1819	7.37	1861	7.85
1778	8.54	1820	7.62	1862	7.56
1779	8.98	1821	8.09	1863	8.11
1780	9.43	1822	8.19	1864	7.98
1781	8.1	1823	7.72	1865	8.18
1782	7.9	1824	8.55	1866	8.29
1783	7.68	1825	8.39	1867	8.44
1784	7.86	1826	8.36	1868	8.25
1785	7.36	1827	8.81	1869	8.43
1786	8.26	1828	8.17	1870	8.2
1787	8.03	1829	7.94	1871	8.12
1788	8.45	1830	8.52	1872	8.19
1789	8.33	1831	7.64	1873	8.35
1790	7.98	1832	7.45	1874	8.43

1875	7.86	1922	8.41	1969	8.6
1876	8.08	1923	8.42	1970	8.7
1877	8.54	1924	8.51	1971	8.6
1878	8.83	1925	8.53	1972	8.5
1879	8.17	1926	8.73	1973	8.95
1880	8.12	1927	8.52	1974	8.47
1881	8.27	1928	8.63	1975	8.74
1882	8.13	1929	8.24	1976	8.35
1883	7.98	1930	8.63	1977	8.85
1884	7.77	1931	8.72	1978	8.69
1885	7.92	1932	8.71	1979	8.73
1886	7.95	1933	8.34	1980	8.98
1887	7.91	1934	8.63	1981	9.17
1888	8.09	1935	8.52	1982	8.64
1889	8.32	1936	8.55	1983	9.03
1890	7.97	1937	8.7	1984	8.69
1891	8.02	1938	8.86	1985	8.66
1892	8.07	1939	8.76	1986	8.83
1893	8.06	1940	8.76	1987	8.99
1894	8.16	1941	8.77	1988	9.2
1895	8.15	1942	8.73	1989	8.92
1896	8.21	1943	8.76	1990	9.23
1897	8.29	1944	8.85	1991	9.18
1898	8.18	1945	8.58	1992	8.84
1899	8.4	1946	8.68	1993	8.87
1900	8.5	1947	8.8	1994	9.04
1901	8.54	1948	8.75	1995	9.35
1902	8.3	1949	8.59	1996	9.04
1903	8.22	1950	8.37	1997	9.2
1904	8.09	1951	8.63	1998	9.52
1905	8.23	1952	8.64	1999	9.29
1906	8.38	1953	8.87	2000	9.2
1907	7.95	1954	8.56	2001	9.41
1908	8.19	1955	8.63	2002	9.57
1909	8.18	1956	8.28	2003	9.53
1910	8.22	1957	8.73	2004	9.32
1911	8.18	1958	8.77	2005	9.7
1912	8.17	1959	8.73	2006	9.53
1913	8.3	1960	8.58	2007	9.73
1914	8.59	1961	8.8	2008	9.43
1915	8.59	1962	8.75	2009	9.51
1916	8.23	1963	8.86	2010	9.7
1917	8.02	1964	8.41	2011	9.52
1918	8.13	1965	8.53	2012	9.51
1919	8.38	1966	8.6	2013	9.61
1920	8.36	1967	8.7	2014	9.57
1921	8.57	1968	8.52	2015	9.83