Data Analyst Nanodegree Program

Project: Explore Weather Trends

Udacity

Project: Explore Weather Trends

August 2018

Data Analyst Nanodegree Program

Project: Explore Weather Trends

Prepared for Udacity

Name	Lee Hong Chen
Username	leehongchen@gmail.com

Submission Date	25 August 2018

UDACITY

CONTENTS

FIGURES

APPENDICES

1 Introduction

1.1 Objective

The objective of this project is analysing local (Singapore) and global temperature data. A comparison will be done for the two data trends of Singapore and global temperature data.

1.2 Scope

The scope of this report focuses on the two datasets which are Singapore and global temperature data. The followings are steps involved in this project:

1. Data extraction from temperature database by using SQL
2. Accessing and analysing data in csv file by using Excel
3. Plotting time series plot (line chart) by Python – Pandas, Matplotlib
4. Data observation

2 Data Extraction

In this Section, Singapore and global temperature data were extracted from database, which provided by Udacity Nanodegree Program. The data extraction was carried out by SQL query below:

```
SELECT gd.year, gd.avg_temp as global_avg_temp, cd.avg_temp as singapore_avg_temp
FROM global_data gd
JOIN city_data cd
ON gd.year=cd.year
WHERE cd.city='Singapore'
```

Figure 2-1 shows that temperature data of Singapore and global were extracted from the database by using SQL query above.

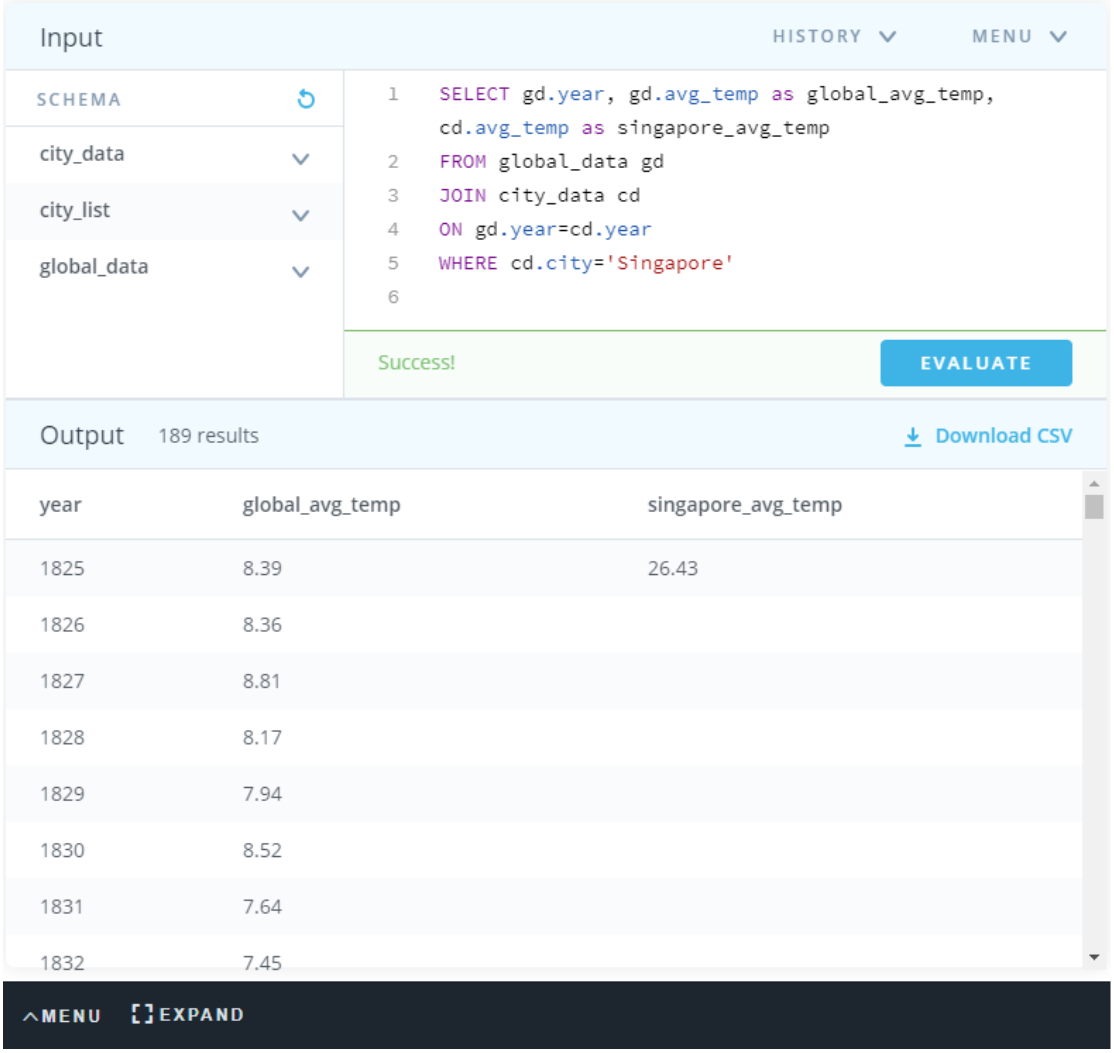

Figure 2-1 Data extraction from the temperature database by using SQL workspace

The extracted temperature data of city (Singapre) and global was exported from database to csv file and saved as “Temperature_Raw_Data.csv”. All extracted data are shown at Appendix A

3 Accessing and Analysing Temperature Data

After data extraction, downloaded csv file (Temperature_Raw_Data.csv) was accessed by using Microsoft Excel. Furthermore, 10 years moving average were calculated for both datasets (Singapore and Global Temperature Data) by using the excel formula below, as shown at Figure 3-1:

"=Average(Cell01:Cell10)"

	A	B	C	D	E
1	Year	Global_avg_temp	10yrs Moving Average of Global Temperature	Singapore_avg_temp	10yrs Moving Average of Singapore Temperature
2	1825	8.39		26.43	
3	1826	8.36			
4	1827	8.81			
5	1828	8.17			
6	1829	7.94			
7	1830	8.52			
8	1831	7.64			
9	1832	7.45			
10	1833	8.01			
11	1834	8.15	=AVERAGE(B2:B11)		26.43
12	1835	7.39	8.04		
13	1836	7.70	7.98		
14	1837	7.38	7.84		
15	1838	7.51	7.77		
16	1839	7.63	7.74	25.79	25.79
17	1840	7.80	7.67	25.89	25.84
18	1841	7.69	7.67	25.98	25.89
19	1842	8.02	7.73	26.14	25.95

Figure 3-1 10 years Moving Average of Singapore and Global Temperature Data

After moving average calculation, the excel file was saved as "Moving_Average_Temperature_Data.csv". All 10 years moving average temperature data are shown at Appendix B.

The following are key considerations of the data trend visualization:

- **10 Years Moving Average (Smoothing method)** were used to smooth temperature data to make it easier to observe long term trends and not get lost in yearly fluctuations. Moving Average is one of smoothing methods (Filtering), which is a procedure to remove variation revealing more clearly the underlying trend and cyclic components in a set of time series data
- **Plotting Line Chart** was considered to present moving average temperature data with x-axis of year and y-axis of temperature levels, as presented at Section 4

4 Plotting Time Series

As mentioned at Section 3, 10 years moving average temperature data of Singapore and Global were presented in line chart in order to visualize Singapore and Global temperature trends. All line charts were plotted by Python–Matplotlib and Python-Pandas.

With the purpose of make observations about the similarities and differences between the world avarages and singapore's avarages, as well as overall trends, the following line charts were plotted, as presented at Figure 5-1 to Figure 5-3:

- 10 Years Moving Average Temperature Data of Singapore
- 10 Years Moving Average Temperature Data of Global
- 10 Years Moving Average Temperature Data of Singapore and Global

The following are the python code used for preparing three plot above. The python code was executed in PyCharm.

```
import pandas as pd
import matplotlib.pyplot as plt

# Read Temperature Data from CSV file
df = pd.read_csv('Input\\Moving_Average_Temperature_Data.csv')

#  Plotting 10 years Moving Average Temperature of Singapore and Global
plt.plot(df['Year'],
df['10yrs_Moving_Average_of_Global_Temperature'],label='Global')
plt.plot(df['Year'],
df['10yrs_Moving_Average_of_Singapore_Temperature'],label='Singapore',
color='red')

plt.title('10 Years Moving Average Temperature Data of Singapore and Global',
fontsize=25)
plt.xlabel('Year',fontsize=20)
plt.ylabel('Temperature, Celsius',fontsize=20)

plt.legend(loc='upper left',fontsize=15)
plt.show()

# Plotting 10 years Moving Average Temperature Temperature Data of Singapore
plt.plot(df['Year'],
df['10yrs_Moving_Average_of_Singapore_Temperature'],label='Singapore',
color='red')

plt.title('10 Years Moving Average Temperature Data of Singapore',
fontsize=25)
plt.xlabel('Year',fontsize=20)
plt.ylabel('Temperature, Celsius',fontsize=20)

plt.legend(loc='upper left',fontsize=15)
```

```
plt.show()

# Plotting 10 years Moving Average Temperature Temperature Data of Global
plt.plot(df['Year'],
df['10yrs_Moving_Average_of_Global_Temperature'],label='Global')

plt.title('10 Years Moving Average Temperature Data of Global', fontsize=25)
plt.xlabel('Year',fontsize=20)
plt.ylabel('Temperature, Celsius',fontsize=20)

plt.legend(loc='upper left', fontsize=15)

plt.show()
```

5 Data Observation

Our world is getting warmer. This is caused by increasing quantity of greenhouse gases in the our atmosphere released from human activities, like deforestation and the burning of fossil fuels.

From Figure 5-1 to Figure 5-3, we could have the following observations on Singapore and Global temperature data (10 years Moving Average) recorded from 1825 to 2013:

1. Singapore temperature is significantly higher than Global during this monitoring period, as presented at Figure 5-1. This difference is consistent over 188 years (1825-2013).
2. Figure 5-2 and Figure 5-3 show that Singapore and Global temperature are progressively increasing over this monitoring period.
3. Singapore and Global temperature changes recorded during this monitoring period are generally comparable. Both changes are approximately 1.5 celcius.
4. Based on all line charts below, we could notice that overall temperature trend of world is progressively increasing. Therefore, both datasets provide conclusive proof that the world is getting hotter. In general, overall world temperature increasing trend is consistent over 188 years (from 1825 to 2013)

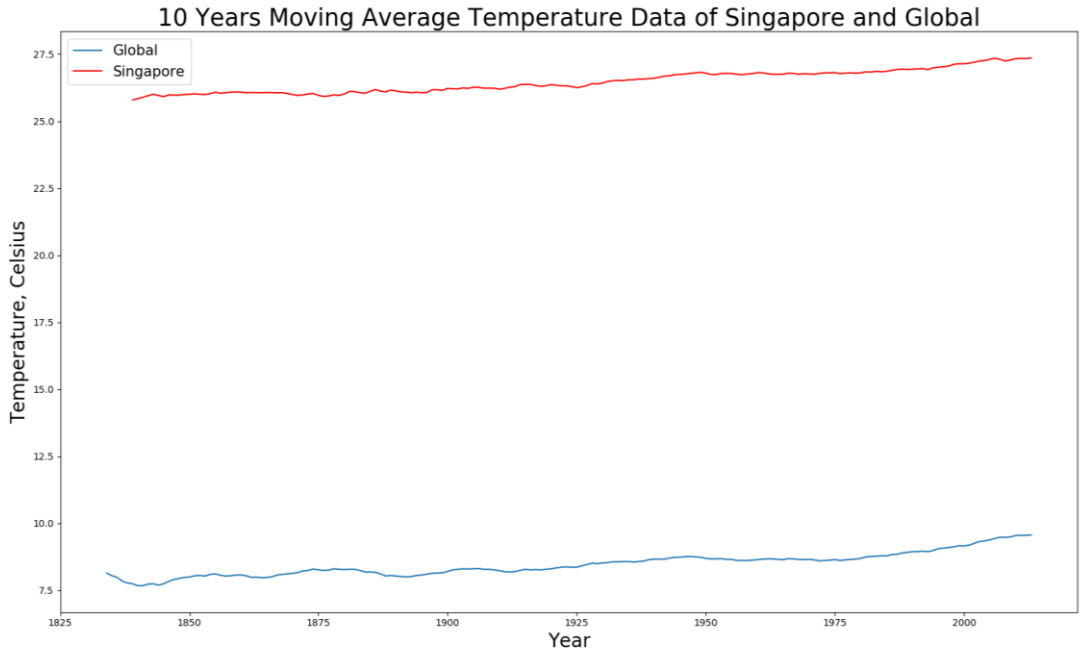

Figure 5-1 10 Years Moving Average Temperature Data of Singapore and Global

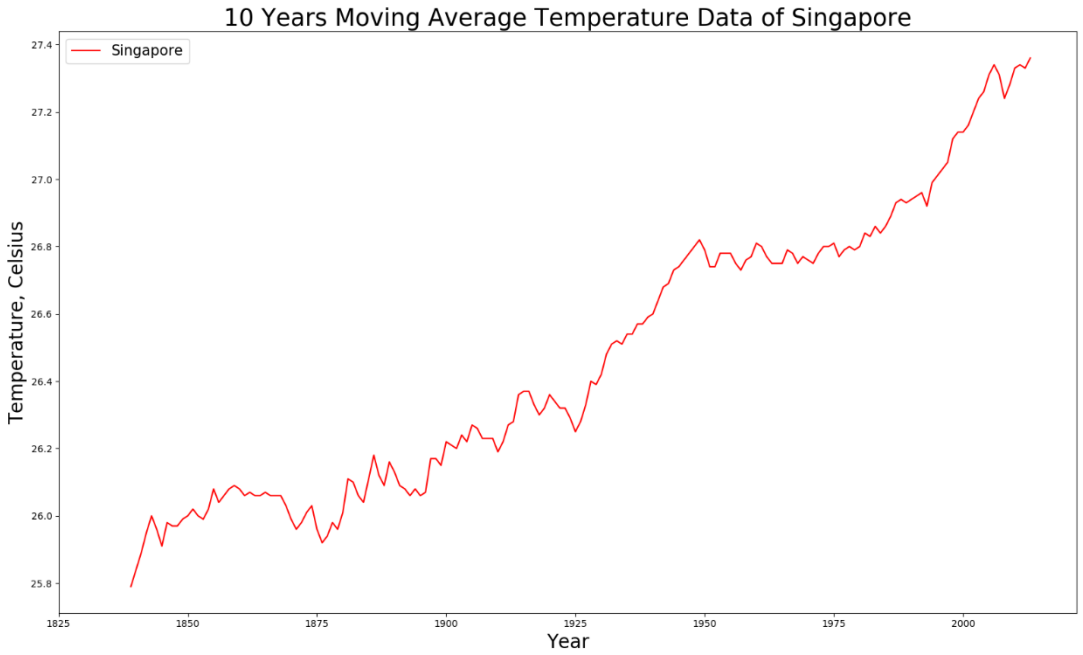

Figure 5-2 10 Years Moving Average Temperature Data of Singapore

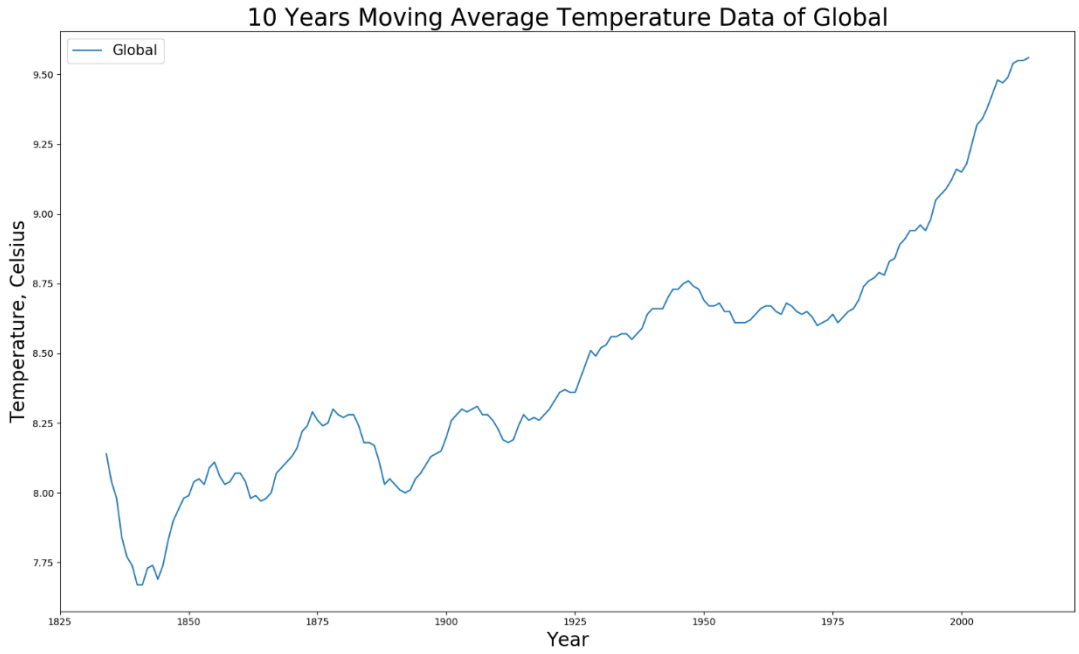

Figure 5-3 10 Years Moving Average Temperature Data of Global

6 Conclusions

Based on observation above, there is sufficient evidence that temperature have risen during the last few hundred years. Due to high temperature, heat waves will arises at whole world. Besides that, a chain reaction of other changes around the world are likely to happens, due to higher temperature.

Therefore, we shall make big changes for this problem. For example, we can substitute fossil fuels needed for our daily energy uses by using renewable resources.

Appendices

Appendix A: Singapore and Global Temperature Data

year	global_avg_temp	singapore_avg_temp
1825	8.39	26.43
1826	8.36	
1827	8.81	
1828	8.17	
1829	7.94	
1830	8.52	
1831	7.64	
1832	7.45	
1833	8.01	
1834	8.15	
1835	7.39	
1836	7.70	
1837	7.38	
1838	7.51	
1839	7.63	25.79
1840	7.80	25.89
1841	7.69	25.98
1842	8.02	26.14
1843	8.17	26.22
1844	7.65	25.73
1845	7.85	25.62
1846	8.55	26.45
1847	8.09	25.88
1848	7.98	
1849	7.98	
1850	7.90	26.01
1851	8.18	26.09
1852	8.10	25.97
1853	8.04	26.18
1854	8.21	25.95
1855	8.11	26.08
1856	8.00	26.17
1857	7.76	
1858	8.10	26.15
1859	8.25	26.24
1860	7.96	25.93
1861	7.85	25.89

1862	7.56	
1863	8.11	26.09
1864	7.98	25.92
1865	8.18	26.15
1866	8.29	26.12
1867	8.44	26.04
1868	8.25	26.15
1869	8.43	25.98
1870	8.20	25.57
1871	8.12	25.58
1872	8.19	26.17
1873	8.35	26.40
1874	8.43	26.09
1875	7.86	25.47
1876	8.08	25.72
1877	8.54	26.25
1878	8.83	26.57
1879	8.17	25.75
1880	8.12	26.12
1881	8.27	26.55
1882	8.13	26.10
1883	7.98	26.00
1884	7.77	25.86
1885	7.92	26.22
1886	7.95	26.40
1887	7.91	25.60
1888	8.09	26.26
1889	8.32	26.46
1890	7.97	25.84
1891	8.02	26.14
1892	8.07	26.04
1893	8.06	25.80
1894	8.16	25.99
1895	8.15	26.09
1896	8.21	26.46
1897	8.29	26.58
1898	8.18	26.30
1899	8.40	26.22
1900	8.50	26.53
1901	8.54	26.06
1902	8.30	25.98
1903	8.22	26.15

1904	8.09	25.85
1905	8.23	26.61
1906	8.38	26.33
1907	7.95	26.30
1908	8.19	26.22
1909	8.18	26.29
1910	8.22	26.11
1911	8.18	26.35
1912	8.17	26.52
1913	8.30	26.26
1914	8.59	26.65
1915	8.59	26.68
1916	8.23	26.29
1917	8.02	25.92
1918	8.13	25.90
1919	8.38	26.55
1920	8.36	26.44
1921	8.57	26.23
1922	8.41	26.28
1923	8.42	26.26
1924	8.51	26.38
1925	8.53	26.22
1926	8.73	26.60
1927	8.52	26.45
1928	8.63	26.58
1929	8.24	26.46
1930	8.63	26.72
1931	8.72	26.88
1932	8.71	26.53
1933	8.34	26.37
1934	8.63	26.25
1935	8.52	26.55
1936	8.55	26.60
1937	8.70	26.73
1938	8.86	26.64
1939	8.76	26.58
1940	8.76	26.86
1941	8.77	27.28
1942	8.73	26.92
1943	8.76	26.49
1944	8.85	26.66
1945	8.58	26.64

1946	8.68	26.83
1947	8.80	26.93
1948	8.75	26.80
1949	8.59	26.74
1950	8.37	26.56
1951	8.63	26.84
1952	8.64	26.89
1953	8.87	26.93
1954	8.56	26.65
1955	8.63	26.64
1956	8.28	26.47
1957	8.73	26.79
1958	8.77	27.08
1959	8.73	26.88
1960	8.58	26.93
1961	8.80	26.73
1962	8.75	26.56
1963	8.86	26.77
1964	8.41	26.66
1965	8.53	26.63
1966	8.60	26.89
1967	8.70	26.68
1968	8.52	26.75
1969	8.60	27.06
1970	8.70	26.85
1971	8.60	26.61
1972	8.50	26.94
1973	8.95	26.96
1974	8.47	26.67
1975	8.74	26.64
1976	8.35	26.56
1977	8.85	26.82
1978	8.69	26.90
1979	8.73	26.96
1980	8.98	26.96
1981	9.17	26.98
1982	8.64	26.84
1983	9.03	27.30
1984	8.69	26.48
1985	8.66	26.82
1986	8.83	26.85
1987	8.99	27.18

1988	9.20	27.07
1989	8.92	26.79
1990	9.23	27.13
1991	9.18	26.99
1992	8.84	27.00
1993	8.87	26.93
1994	9.04	27.09
1995	9.35	27.07
1996	9.04	27.04
1997	9.20	27.39
1998	9.52	27.78
1999	9.29	27.02
2000	9.20	27.10
2001	9.41	27.19
2002	9.57	27.41
2003	9.53	27.31
2004	9.32	27.33
2005	9.70	27.49
2006	9.53	27.33
2007	9.73	27.13
2008	9.43	27.07
2009	9.51	27.47
2010	9.70	27.60
2011	9.52	27.28
2012	9.51	27.30
2013	9.61	27.60

UDACITY

Appendix B: 10 years Moving Average Temperature Data for Singapore and Global

year	global_avg_temp	10yrs_Moving_Average_of_Global Temperature	singapore_avg_temp	10yrs_Moving_Average_of_Singapore_Temperature
1825	8.39	NA	26.43	NA
1826	8.36	NA	NA	NA
1827	8.81	NA	NA	NA
1828	8.17	NA	NA	NA
1829	7.94	NA	NA	NA
1830	8.52	NA	NA	NA
1831	7.64	NA	NA	NA
1832	7.45	NA	NA	NA
1833	8.01	NA	NA	NA
1834	8.15	8.14	NA	26.43
1835	7.39	8.04	NA	NA
1836	7.70	7.98	NA	NA
1837	7.38	7.84	NA	NA
1838	7.51	7.77	NA	NA
1839	7.63	7.74	25.79	25.79
1840	7.80	7.67	25.89	25.84
1841	7.69	7.67	25.98	25.89
1842	8.02	7.73	26.14	25.95
1843	8.17	7.74	26.22	26.00
1844	7.65	7.69	25.73	25.96
1845	7.85	7.74	25.62	25.91
1846	8.55	7.83	26.45	25.98
1847	8.09	7.90	25.88	25.97
1848	7.98	7.94	NA	25.97
1849	7.98	7.98	NA	25.99
1850	7.90	7.99	26.01	26.00
1851	8.18	8.04	26.09	26.02
1852	8.10	8.05	25.97	26.00
1853	8.04	8.03	26.18	25.99
1854	8.21	8.09	25.95	26.02
1855	8.11	8.11	26.08	26.08
1856	8.00	8.06	26.17	26.04
1857	7.76	8.03	NA	26.06
1858	8.10	8.04	26.15	26.08
1859	8.25	8.07	26.24	26.09
1860	7.96	8.07	25.93	26.08

1861	7.85	8.04	25.89	26.06
1862	7.56	7.98	NA	26.07
1863	8.11	7.99	26.09	26.06
1864	7.98	7.97	25.92	26.06
1865	8.18	7.98	26.15	26.07
1866	8.29	8.00	26.12	26.06
1867	8.44	8.07	26.04	26.06
1868	8.25	8.09	26.15	26.06
1869	8.43	8.11	25.98	26.03
1870	8.20	8.13	25.57	25.99
1871	8.12	8.16	25.58	25.96
1872	8.19	8.22	26.17	25.98
1873	8.35	8.24	26.40	26.01
1874	8.43	8.29	26.09	26.03
1875	7.86	8.26	25.47	25.96
1876	8.08	8.24	25.72	25.92
1877	8.54	8.25	26.25	25.94
1878	8.83	8.30	26.57	25.98
1879	8.17	8.28	25.75	25.96
1880	8.12	8.27	26.12	26.01
1881	8.27	8.28	26.55	26.11
1882	8.13	8.28	26.10	26.10
1883	7.98	8.24	26.00	26.06
1884	7.77	8.18	25.86	26.04
1885	7.92	8.18	26.22	26.11
1886	7.95	8.17	26.40	26.18
1887	7.91	8.11	25.60	26.12
1888	8.09	8.03	26.26	26.09
1889	8.32	8.05	26.46	26.16
1890	7.97	8.03	25.84	26.13
1891	8.02	8.01	26.14	26.09
1892	8.07	8.00	26.04	26.08
1893	8.06	8.01	25.80	26.06
1894	8.16	8.05	25.99	26.08
1895	8.15	8.07	26.09	26.06
1896	8.21	8.10	26.46	26.07
1897	8.29	8.13	26.58	26.17
1898	8.18	8.14	26.30	26.17
1899	8.40	8.15	26.22	26.15
1900	8.50	8.20	26.53	26.22
1901	8.54	8.26	26.06	26.21
1902	8.30	8.28	25.98	26.20

UDACITY

1903	8.22	8.30	26.15	26.24
1904	8.09	8.29	25.85	26.22
1905	8.23	8.30	26.61	26.27
1906	8.38	8.31	26.33	26.26
1907	7.95	8.28	26.30	26.23
1908	8.19	8.28	26.22	26.23
1909	8.18	8.26	26.29	26.23
1910	8.22	8.23	26.11	26.19
1911	8.18	8.19	26.35	26.22
1912	8.17	8.18	26.52	26.27
1913	8.30	8.19	26.26	26.28
1914	8.59	8.24	26.65	26.36
1915	8.59	8.28	26.68	26.37
1916	8.23	8.26	26.29	26.37
1917	8.02	8.27	25.92	26.33
1918	8.13	8.26	25.90	26.30
1919	8.38	8.28	26.55	26.32
1920	8.36	8.30	26.44	26.36
1921	8.57	8.33	26.23	26.34
1922	8.41	8.36	26.28	26.32
1923	8.42	8.37	26.26	26.32
1924	8.51	8.36	26.38	26.29
1925	8.53	8.36	26.22	26.25
1926	8.73	8.41	26.60	26.28
1927	8.52	8.46	26.45	26.33
1928	8.63	8.51	26.58	26.40
1929	8.24	8.49	26.46	26.39
1930	8.63	8.52	26.72	26.42
1931	8.72	8.53	26.88	26.48
1932	8.71	8.56	26.53	26.51
1933	8.34	8.56	26.37	26.52
1934	8.63	8.57	26.25	26.51
1935	8.52	8.57	26.55	26.54
1936	8.55	8.55	26.60	26.54
1937	8.70	8.57	26.73	26.57
1938	8.86	8.59	26.64	26.57
1939	8.76	8.64	26.58	26.59
1940	8.76	8.66	26.86	26.60
1941	8.77	8.66	27.28	26.64
1942	8.73	8.66	26.92	26.68
1943	8.76	8.70	26.49	26.69
1944	8.85	8.73	26.66	26.73

1945	8.58	8.73	26.64	26.74
1946	8.68	8.75	26.83	26.76
1947	8.80	8.76	26.93	26.78
1948	8.75	8.74	26.80	26.80
1949	8.59	8.73	26.74	26.82
1950	8.37	8.69	26.56	26.79
1951	8.63	8.67	26.84	26.74
1952	8.64	8.67	26.89	26.74
1953	8.87	8.68	26.93	26.78
1954	8.56	8.65	26.65	26.78
1955	8.63	8.65	26.64	26.78
1956	8.28	8.61	26.47	26.75
1957	8.73	8.61	26.79	26.73
1958	8.77	8.61	27.08	26.76
1959	8.73	8.62	26.88	26.77
1960	8.58	8.64	26.93	26.81
1961	8.80	8.66	26.73	26.80
1962	8.75	8.67	26.56	26.77
1963	8.86	8.67	26.77	26.75
1964	8.41	8.65	26.66	26.75
1965	8.53	8.64	26.63	26.75
1966	8.60	8.68	26.89	26.79
1967	8.70	8.67	26.68	26.78
1968	8.52	8.65	26.75	26.75
1969	8.60	8.64	27.06	26.77
1970	8.70	8.65	26.85	26.76
1971	8.60	8.63	26.61	26.75
1972	8.50	8.60	26.94	26.78
1973	8.95	8.61	26.96	26.80
1974	8.47	8.62	26.67	26.80
1975	8.74	8.64	26.64	26.81
1976	8.35	8.61	26.56	26.77
1977	8.85	8.63	26.82	26.79
1978	8.69	8.65	26.90	26.80
1979	8.73	8.66	26.96	26.79
1980	8.98	8.69	26.96	26.80
1981	9.17	8.74	26.98	26.84
1982	8.64	8.76	26.84	26.83
1983	9.03	8.77	27.30	26.86
1984	8.69	8.79	26.48	26.84
1985	8.66	8.78	26.82	26.86
1986	8.83	8.83	26.85	26.89

1987	8.99	8.84	27.18	26.93
1988	9.20	8.89	27.07	26.94
1989	8.92	8.91	26.79	26.93
1990	9.23	8.94	27.13	26.94
1991	9.18	8.94	26.99	26.95
1992	8.84	8.96	27.00	26.96
1993	8.87	8.94	26.93	26.92
1994	9.04	8.98	27.09	26.99
1995	9.35	9.05	27.07	27.01
1996	9.04	9.07	27.04	27.03
1997	9.20	9.09	27.39	27.05
1998	9.52	9.12	27.78	27.12
1999	9.29	9.16	27.02	27.14
2000	9.20	9.15	27.10	27.14
2001	9.41	9.18	27.19	27.16
2002	9.57	9.25	27.41	27.20
2003	9.53	9.32	27.31	27.24
2004	9.32	9.34	27.33	27.26
2005	9.70	9.38	27.49	27.31
2006	9.53	9.43	27.33	27.34
2007	9.73	9.48	27.13	27.31
2008	9.43	9.47	27.07	27.24
2009	9.51	9.49	27.47	27.28
2010	9.70	9.54	27.60	27.33
2011	9.52	9.55	27.28	27.34
2012	9.51	9.55	27.30	27.33
2013	9.61	9.56	27.60	27.36