Name: Shukhrat Khuseynov **ID:** 0070495 **Date:** 01.12.2019

Course: DASC 501 **Assignment 3 - Weather Data Analysis**

Discussion (Interpretation)

The "oxforddata.txt" was uploaded from its official site, including the entries until October 2019, since the "weather_data_Oxford.txt" file in BlackBoard is outdated, but the latter will also work properly in my code. The code includes both parts, their output is given below:

Part 1. Average minimum temperature for each month:

Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec
1.49	1.55	2.47	4.42	7.22	10.32	12.28	12.04	9.93	6.95	3.75	2.08

Maximum: 12.28

Average of mean daily minimum temperature for each month in each decade:
1853 - 1859 : 5.56
1860 - 1869 : 5.88
1870 - 1879 : 5.85
1880 - 1889 : 5.68
1890 - 1899 : 5.67
1900 - 1909 : 5.86
1910 - 1919 : 6.03
1920 - 1929 : 5.92
1930 - 1939 : 6.21
1940 - 1949 : 6.16
1950 - 1959 : 6.16
1960 - 1969 : 6.28
1970 - 1979 : 6.36
1980 - 1989 : 6.54
1990 - 1999 : 7.03
2000 - 2009 : 7.19
2010 - 2019 : 7.12

Part 2. Average number of sunshine hours for each month:

Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec
57.63	73.77	116.96	155.16	191.65	198.12	196.62	183.66	139.16	104.68	67.79	52.9

Maximum: 198.12

Average of sunshine hours for each month in each decade:
1929 - 1929 : 145.03
1930 - 1939 : 119.43
1940 - 1949 : 125.34
1950 - 1959 : 130.00
1960 - 1969 : 123.48
1970 - 1979 : 125.56

1980 - 1989 : 124.11
1990 - 1999 : 135.94
2000 - 2009 : 134.01
2010 - 2019 : 135.26

Note: the provisional data was used in both parts, where possible.

The computations are given in the output of both parts. In the first part, best weather (warmest temperature) is encountered in July, while in the second part best weather (most sunshine hours) is encountered in June, which is quite close and related.

Moreover, decade analysis truly demonstrates the change in climate (it is becoming warmer), which is a popular notion in the world today. Especially this can be observed in the first part, where the range of years is wider, the minimum temperature significantly increases throughout the time. Regarding the second part, if to ignore the incomplete decade (1929-1929), on average the data is also consistent with this idea.

Technical side:

(i) I chose Pandas data frames to implement the second part because it is more suitable and easier than Numpy while working with the tables. Pandas' functions are designed for the tabular data analysis. Numpy would take more time and efforts.

(ii) If to compare the nested lists method and Pandas data frames, it was easier to work with Pandas. In the first part, where the nested lists were used, much more loops were required for numerous iterations. It was more difficult to initialize the lists, to write the indices, etc. On the other hand, Pandas data frames, being used in the second part, required less time and efforts. Pandas' functionality was simplifying many tasks.