**Assignment 5**

Kristina Ghahramanyan

American University of Armenia

ENV105 - Numbers Responsibility Environment

Mr. Sean Reynolds

December 7, 2021

**Essay**

Creating a clean and informative dataset about weather data was an extensive process. The main requirement of such a task was getting acquainted with working in excel, which was accomplished throughout the course lectures.

The actual work began with transferring the necessary data into separate sheets while keeping everything concise and tidy. At this stage, it was essential to thoroughly verify that the values made sense and to make sure no missing cells were present. Afterward, some built-in elementary tools were used on the dataset. One such feature was conditional formatting, performed on one column of each month to highlight its values. I chose to use the column with maximum temperatures and selected a color palette to make it intuitive for a reader to understand the data. The darker, redder cells represented higher values, while the lighter, greener ones included lower numbers.

Next, it was time to use various functions for showing the overall maximum, minimum, and average data. Before looking at the functions, it was essential to use named ranges to create distinct tags for the data I would use. This step was significant because it made the whole process much easier and less time-consuming. For example, I was able to name my entire column of maximum temperatures into a word like “MaxJan20”, which was understandable and shorter than “=”January” ! B1:B20”. After naming all my ranges, I manually created a table with all the high maximum values, low minimums, etc. Then, I needed to compare my data to a historical one, which was quickly created by simply finding the values from the internet.

When the functions were written, and the table was done, it was time to make some graphs to see if any trends or similarities could be found in the dataset. The first chart I chose was a line graph used to show the maximum temperature based on recent and historical data. Both lines of the graph were very similar, but the values of recent data happened to be a little lower almost everywhere. The following graph was a bar chart showing the maximum and minimum temperatures for the combined data of 2019 and 2020. After seeing the peaks of the values, I was able to make some interesting observations about the data. The chart shows that the maximum temperatures are reached in May, and only after September do numbers begin to decrease noticeably. Yet, the bars representing minimum temperatures show significant changes throughout those same months, which could mean that possibly, in that five-month period, temperatures during the day do not change as much as they do at night. My last chart was a scatter plot that shows average humidity in August. The choice was made based on noticing that humidity in that month is somewhat random. One might think that due to high temperatures, August will have consistently low humidity results, but in this case, we do not see that picture.

To finish the work with this dataset, I needed to do some calculations for finding standard deviation, standard error and calculate how much of my data is located inside each divided part (68%, 95%, etc.). The results showed that most of my data (about 54%) is inside the 1 sigma area (68% of the whole area). Now all that is left is to make a consolidated sheet containing all my data for both years 2019 and 2020. That step allowed me to generate a pivot table for comparing it to my original, manually created table to make sure everything matches. Creating a pivot table was relatively easy, as I made more than six of them for my group project. Firstly, I had to specify that all the values must be located in the rows and the dates in the columns. Secondly, the values of my data were changed to their respective types (from the automatically assigned sum to maximum, average, and minimum). Lastly, the format of values was changed to “Number,” and the decimals were modified to contain only two digits.

When everything was done, I went through all my sheets and fixed some minor mistakes with cell sizes, value formats, and named ranges. The charts were also relocated and double-checked, as well as the sigma calculations, which had to be redone due to my mistakes in previous steps. Then, the weather data report was finally clean, easy to read, and ready to be submitted.