**Assignment 1 - Report**

# **Introduction:**

A picture containing text, writing implement, stationary, pencil

Description automatically generatedFor this assignment, I chose to research the weather in the UK by using the Mean Central England Temperature (oC)[[1]](#footnote-1) dataset from the Met Office. This dataset contains the mean temperature for each month and the annual temperature from 1659 to 2023 (I have not included the data from 2023 as the data is not complete for the year). I have taken the lowest and highest mean temperature values for each year and added them to new lists. From this, I could use my knowledge of Data Science and the UK weather to analyse the data.

# **Analysis of the Data:**

Maximum Temperatures per Year:

After some basic pre-processing and manipulation, I determined that July is the most common month with the highest temperature, with an average temperature of 16.4oC. The hottest mean temperature recorded was 19.8oC in July 2006.

Chart, histogram

Description automatically generatedFrom this, I constructed a line graph (figure 1) which shows there has been a steady increase in the maximum mean yearly temperatures. However, there were a few dips, for example, between 8-12 and 17-18.5 on the time axis. There are also some extreme/outliers in the data as shown by the spikes in the data (a few are shown by the green circles in figure 1).

A picture containing text, writing implement, stationary, pencil

Description automatically generatedI also constructed a histogram (figure 2) which shows that the data is roughly the shape of a normal distribution, however, it is slightly positively skewed to the right, as the mean is 0.2oC higher than the median, which is backed up further as the skewness value was 0.326oC. The left-hand side tail seems to represent a normal distribution less compared to the right-hand side tail, which is also backed up by a Fisher’s kurtosis value of -0.066oC, meaning the frequency density of the maximum temperatures per year is very close to being a perfect normal distribution.

Minimum Temperatures per Year:

Chart, line chart

Description automatically generatedAgain, after some basic pre-processing and manipulation, I determined that January is the most common month with the lowest temperature, with an average temperature of 2.3oC. The coldest mean temperature recorded was -3.1oC in July 1795.

From this, I constructed a line graph (figure 3) which easily shows that the overall trend of minimum yearly temperatures is rising steadily. However, there are a few dips in the maximum temperatures, for example, between 6-14 and 16.5-18 on the Time axis. There have also been some extreme/outliers in the data as shown by the spikes in the data (a few are shown by the green circles in figure 3).

I also constructed a histogram (figure 4) which shows that the data roughly follows the shape of a normal distribution, however, it is slightly negatively skewed to the left, as the mean is 0.1oC lower than the median. This is backed up further as when I calculated the skewness value it came to be -0.505oC. The right-hand side tail seems to represent a normal distribution less compared to the left-hand side tail. This is also, backed up by a Fisher’s kurtosis value of -0.208oC.

Overall:

When calculating Spearman’s Rank Correlation Coefficient, which came to be 0.1750 with the corresponding p-value of 0.0008 it showed that there are no linear relationships between the maximum and minimum yearly temperatures as the coefficient was quite close to 0.

# **Summary:**

Overall, the Mean Central England Temperature (oC) dataset shows that from 1659 to 2022, there has been a steady increase in the temperatures being recorded. This is mostly due to an increase in the rate of global warming as more and more greenhouse gases are being emitted every day.

1. Data sourced from: <https://www.metoffice.gov.uk/hadobs/hadcet/data/meantemp_monthly_totals.txt> [↑](#footnote-ref-1)