Datasets

By collecting sets of data at set times during the day within the same area, data sets were collected for the plotting of graphs using matlab.

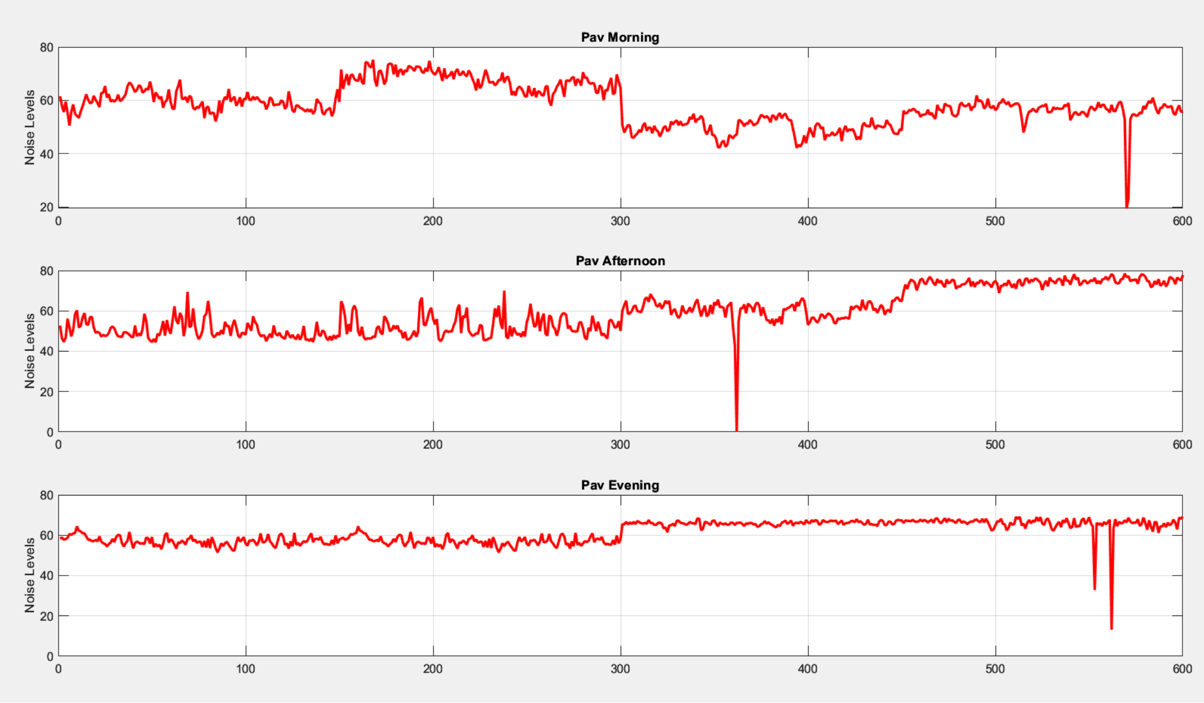


Figure 1. Plots of Pav data sets

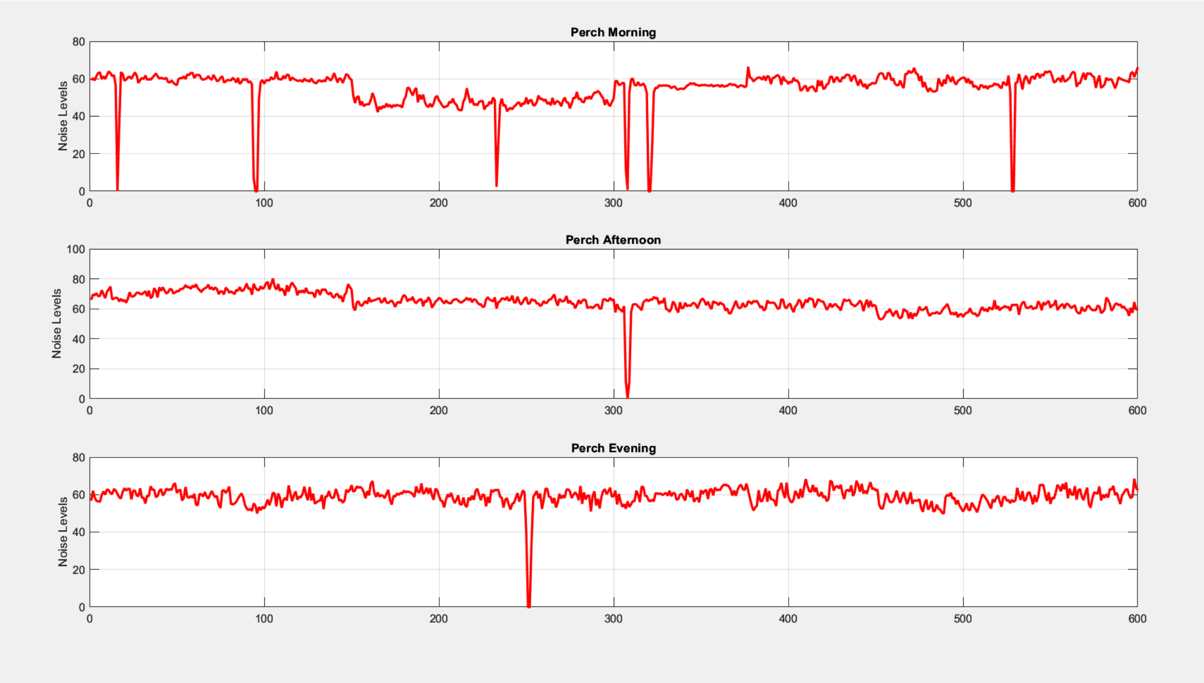


Figure 2. Plots of Perch data sets

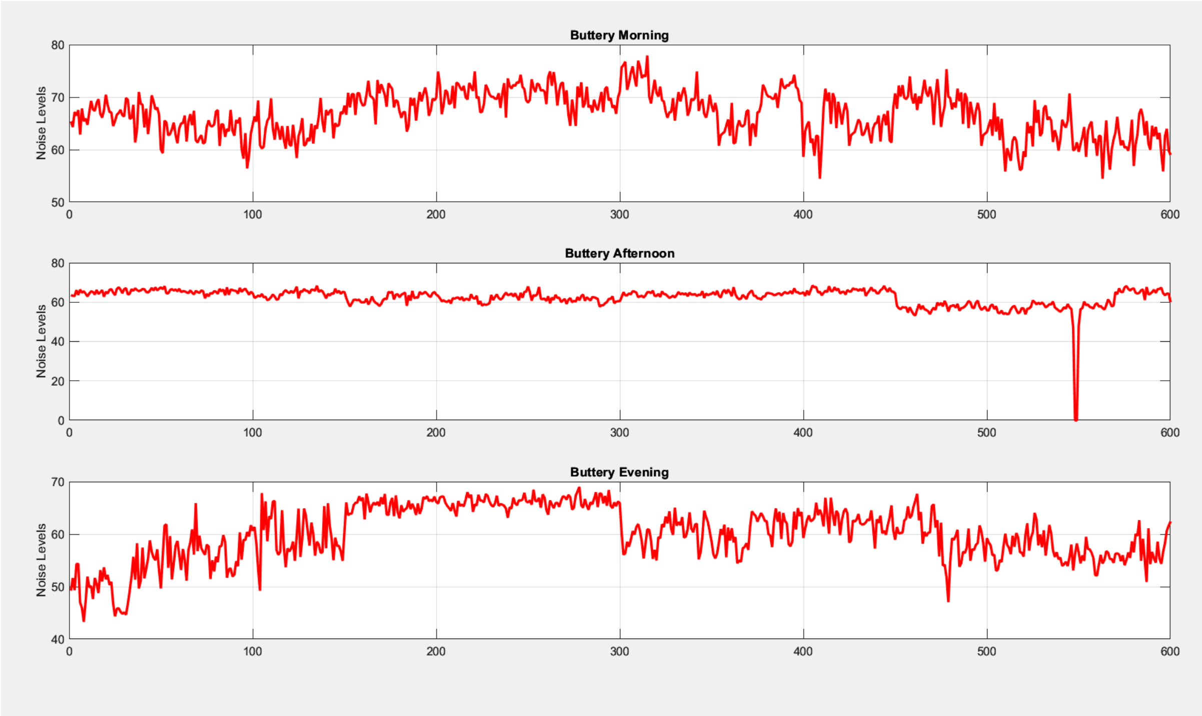


Figure 3. Plots of Buttery data sets

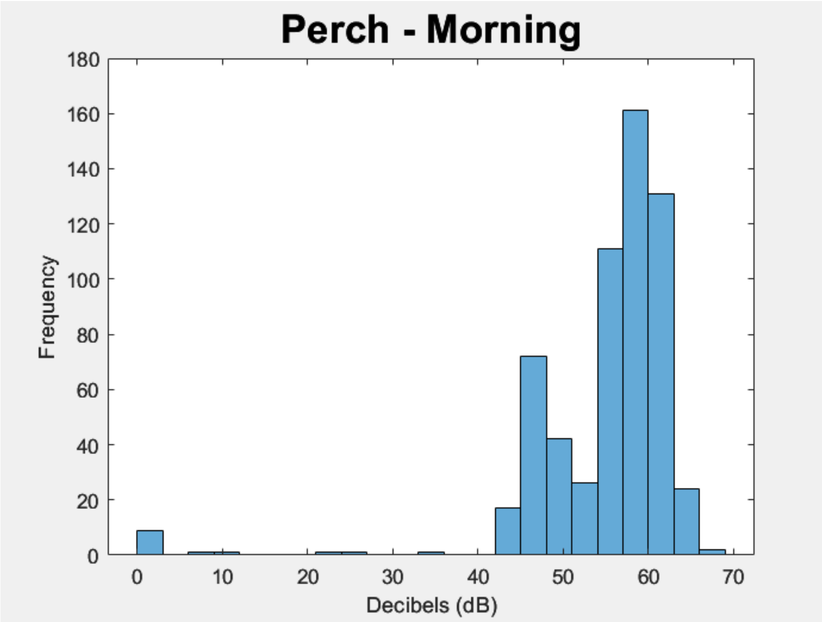
It seems that the Buttery is the noisiest place on average.The high peaks and their frequency indicate that this location sees more higher and frequent noise events compared to The Pav and The Perch. The Pav however is the noisiest place in the evenings which is notably noisier than the other locations.

From looking at the data sets it can also be concluded that the Pav tends to be noisiest in the evenings while the Buttery tends to be noisiest during the afternoons with a similar drop off in sound levels during the mornings and evenings. Both the pav and the buttery exhibit fairly significant changes in sound levels depending on the time of day while the perch seems to remain fairly consistent regardless of time of day. A possible explanation for this phenomena is the fact that the perch is the only place with lecture halls in its immediate vicinity. Both the butter and the pav are removed from lecture halls.

It should also be noted that the recordings are done in intervals of 150 which can explain the shape of the graphs which tend to have a pattern change with every 150 data points. Another point to note about the data is that there are points at which the sound being recorded is at 0 decibels which is highly unrealistic if not impossible without a specialised environment of sound designed muffling and absorption, this shows that the recording device or software used is not entirely accurate.

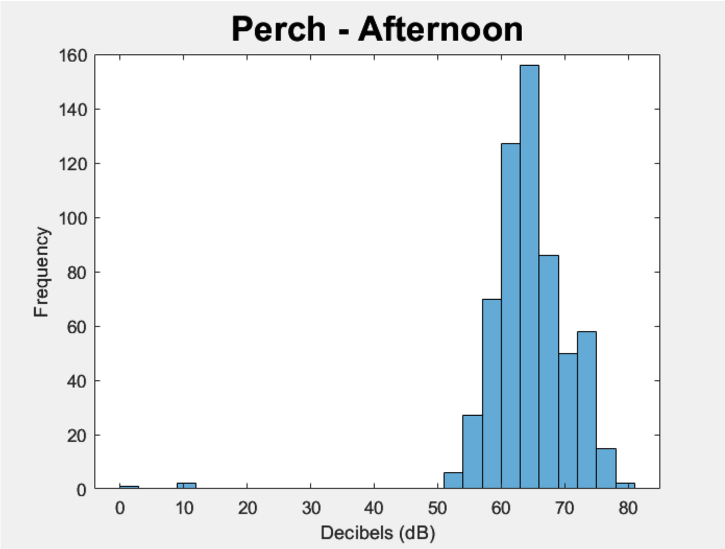
Histograms

Using the MATLAB function histogram we plotted a histogram for each dataset after combining all of the 5 seconds averages for each place of data collection and time frame.

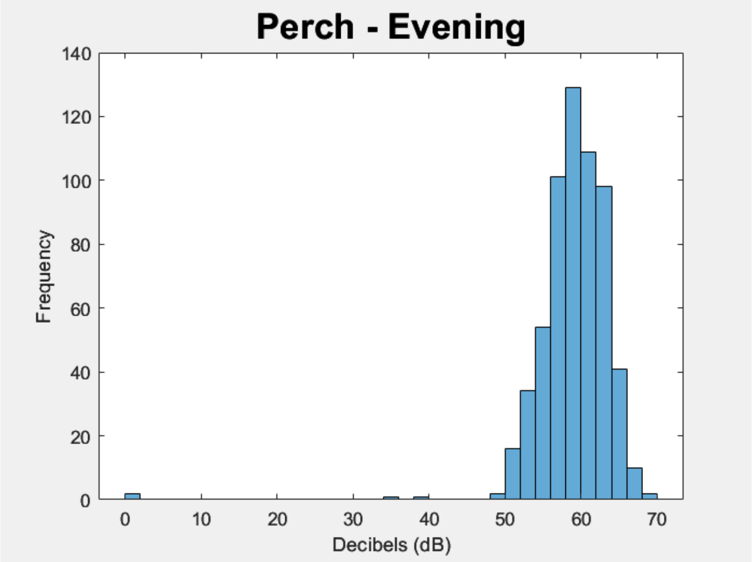


The histogram representing the sound levels for the Perch café in the morning shows data that follows a symmetrical shape with a peak in the centre of the data at 55 - 60 dB. The adjacent values have lower frequencies and make it so the data has a bell shape. Furthermore, as the data is skewed to the right, the mode is less than the median, which is less than the mean. In addition, on the quieter side of the data there is a small peak at 45 dB which leads to the idea that during the data collection, there was an increase in average noise which caused the decibels to shift upwards and create a higher peak to the right of the graph.

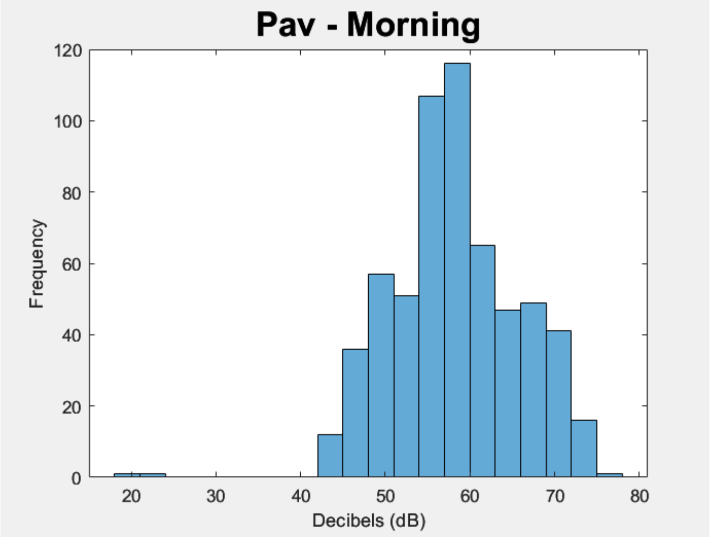
The data also indicates some values recorded at abnormal volumes. Due to the uncertainty and lack of precision of the measuring device it can be considered an outlier.



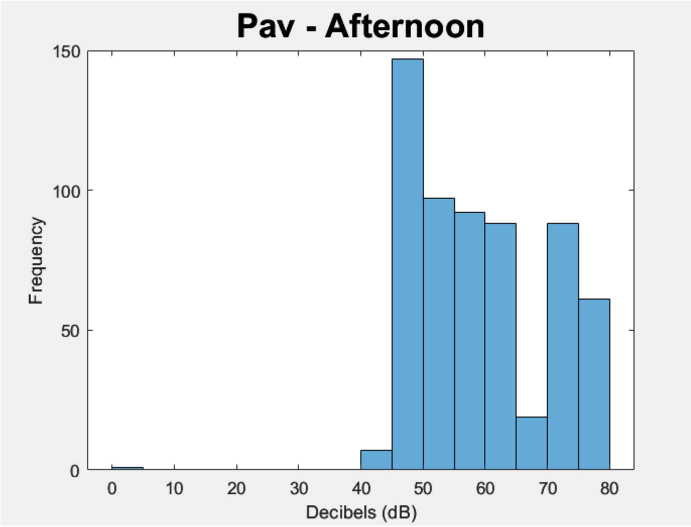
The histogram representing the sound levels for the Perch café in the afternoon shows data that follows a symmetrical shape with a peak in the centre of the data at around 65 dB. The adjacent values have lower frequencies which give the histogram a bell shape. Furthermore, this data is very slightly skewed to the left, indicating that the mean is less than the median, which is less than the mode.The data also indicates some values recorded at abnormal volumes at 0 and 10 dB. Due to the uncertainty and lack of precision of the measuring device, they can be considered outliers.



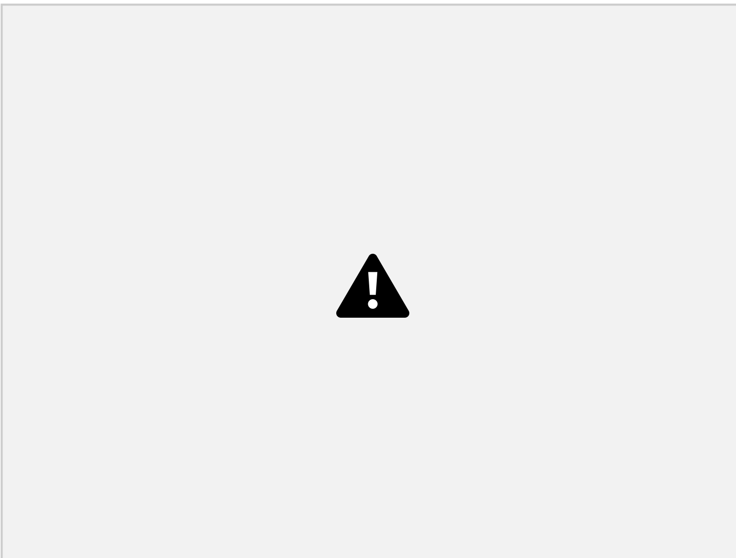
The histogram representing the sound levels for the Perch café in the evenings shows data that follows a symmetrical shape with a peak in the centre of the data at 60 dB. The adjacent values have lower frequencies and make it so the data has a bell shape.
The data also indicates some values recorded at the abnormal volumes of 0, 35 and 40 dB. Due to the uncertainty and lack of precision of the measuring device they can be considered outliers.



The histogram representing the sound levels for the Pav in the morning shows data that follows a symmetrical shape with a peak in the centre of the data at 55 - 60 dB. The adjacent values have lower frequencies and make it so the data has a bell shape. In addition, the peak for this data set has a larger volume range than the others and a bigger drop in frequency with adjacent volumes, indicating louder volume increases in the environment.
The data also indicates some values recorded at the unexpected volumes of around 20 dB. Due to the uncertainty and lack of precision of the measuring device they can be considered outliers.

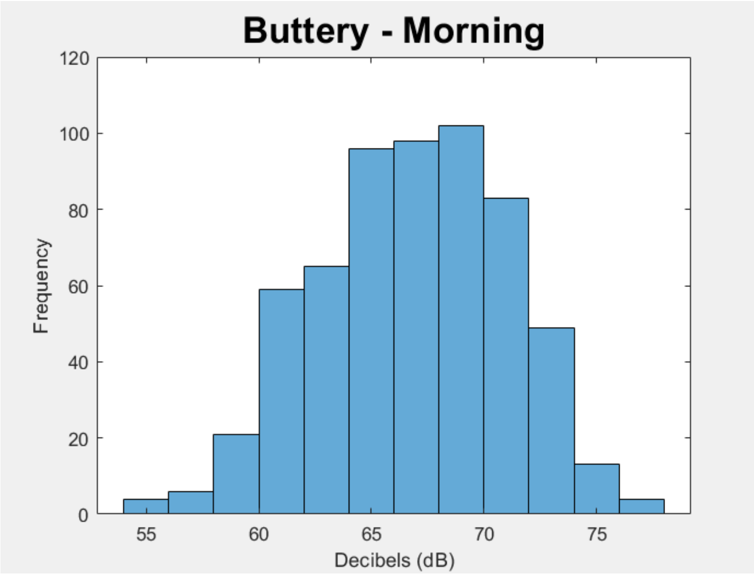


The histogram representing the sound levels for the Pav in the afternoon shows data with a very different shape to the other plots. A single peak can be seen to the left of the data at 50 dB with very few values lower than that. After the peak there is a very constant frequency of volumes with a slight decreasing tendency. This indicates that the volume was fairly constant at around 45 to 50 dB and then the environment got louder and larger values started to appear at a very constant rate.
The data also indicates some values recorded at 0 dB. Due to the uncertainty and lack of precision of the measuring device it can be considered an outlier.

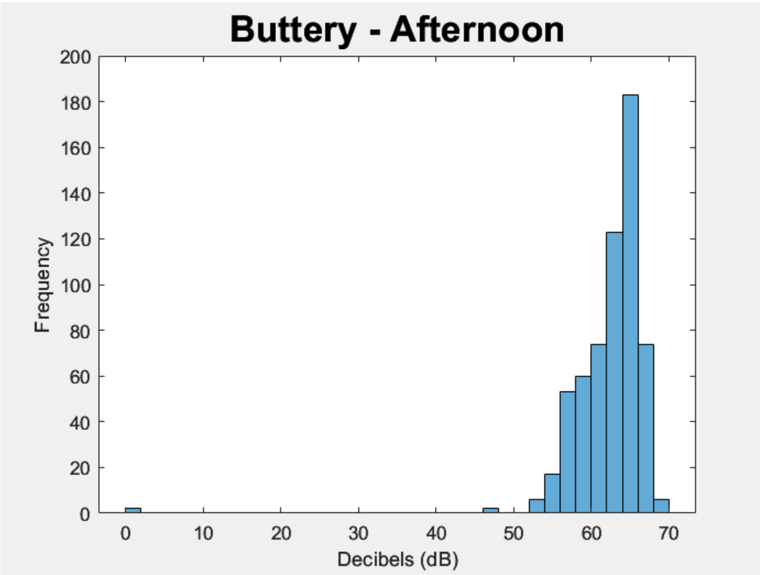


The histogram representing the sound levels for the Pav in the evening shows data very similar to the previous histogram. This one has its peak to the right of the data at 64 - 68 dB and next to the peak in the lower volume side, values with a much lower yet constant frequency except for a slight increase at 56 dB. Similarly, this indicates that the volume was fairly constant at around 64 - 68 dB and then the environment got quieter and smaller values started to appear at a very constant rate.

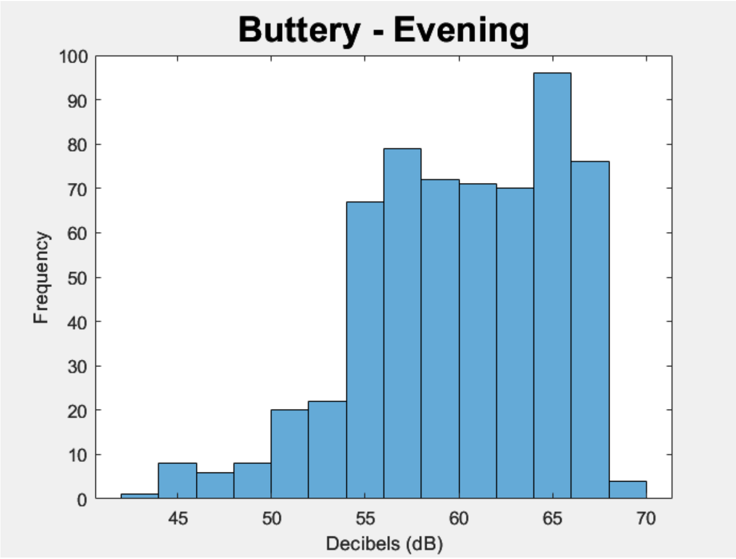
The data also shows a very small amount of values recorded at abnormal volumes (14 - 34 dB). Due to the uncertainty and lack of precision of the measuring device it can be considered an outlier.



The histogram representing the sound levels for the Buttery in the morning shows data that follows a symmetrical shape with a very wide peak in the centre of the data at 65 - 70 dB. The adjacent values have lower frequencies and make it so the data has a bell shape. Furthermore, the data has very little skewness which means that the mean median and mode are all very close to each other but not exactly the same as the histogram is not perfectly symmetrical.
This data does not show significant outliers because by following the shape of the histogram, the position and frequency of the values follow the trend.



The histogram representing the sound levels for the Buttery in the afternoon shows data that follows a symmetrical shape with a peak at 65 dB. The adjacent values have lower frequencies and make it so the data has a skewed bell shape. The skewness of the data is to the right. This indicates that the mode is less than the median, which is less than the mean.
The data also indicates some values recorded at 0 dB. Due to the uncertainty and lack of precision of the measuring device it can be considered an outlier.



The histogram representing the sound levels for the Buttery in the evening shows a shape similar to the one in the morning. In this case there is a constant frequency for the range of values between 55 - 68 dB with a slight peak at 65 dB. The plot then shows values with much lower frequencies at lower volume levels (45 - 55) with a slight increasing tendency as the volume increases. This shape could be due to a constant volume in the environment and then a sudden change that remains constant for the rest of the data collection.

The data also indicates some values recorded at abnormal volumes of around 20 dB. Due to the uncertainty and lack of precision of the measuring device it can be considered an outlier.

Measures of Central Tendencies

	Mean	Median
Perch Morning	55.41833182	57.58550655
Perch Lunch	64.45024412	64.2819079
Perch Evening	58.97634606	59.3385
Buttery Morning	66.8182042	67.163
Buttery Lunch	62.4131538	63.57808182
Buttery Evening	60.09415	60.38
Pav Morning	58.22077893	57.80668637
Pav Lunch	59.25489377	57.48663455

Pav Evening	63.74853712	65.67510429

The central tendency of a data set describes the midpoint of a distribution of values in a dataset. The mean represents the average of all the values and the median represents the middle number in a ordinally categorised data set. The mode represents the variable that appears most often within a data set; however, in this case, the numbers vary for each measurement so therefore every variable is unique, and the mode is not applicable.

It is important to consider the mean and median values above to allow for a comprehensive analysis to be made on the various areas. The mean values offer a quantitative summary of the central tendencies in the various area's acoustic environments. The values represent the average noise levels over the different days that the measurements were obtained. From the table above, it can be observed that the mean gain values for the Perch ranged from 55 to 59 which suggests that there is a highly comparable central tendency between each of the time periods. Similarly, for the buttery and the Pav, the overall mean values range from 60 to 66 and 58 to 64 respectively.

The median values offer insight into the distribution of the noise levels over the entire dataset. This measurement is not affected by any outliers and captures the intensity of the noise levels that are very common to the different areas. The median values overall in all areas range from 57 to 67 which indicates similar noise concentrations in each level.

The mean and the median both offer important insights into the overall data distribution, and both measurements are hugely important when analysing different data sets.

Trimmed Mean

Place	5% trimmed mean	20% trimmed mean
Buttery morning	66.8348	66.8432
Buttery afternoon	62.5846	62.5236
Buttery evening	60.0966	60.1163
Pav morning	58.3521	58.3811
Pav afternoon	59.2623	59.1189
Pav evening	63.6555	63.8439
Perch morning	55.8948	55.9828
Perch afternoon	64.7096	64.7677
Perch evening	59.2166	59.2979

Using the 5% and 20% trimmed mean values is a useful way of reducing the impact of outliers on measures of central tendency. The trimmed means are considered more robust and reliable because they take into account only the central portion of the data and ignore the extreme values that might skew the results. Interestingly, the trimmed means generally align with the standard mean values, indicating that the effect of the outliers on the overall central tendency is not as significant as previously thought.

Conclusion

Based on the analysis provided in parts II to IV, it can be concluded that the Buttery generally has higher noise levels than the Pav and the Perch, especially during the morning and lunch times. This is supported by the consistently higher mean, median, and trimmed mean values of the Buttery when compared to the Pav and the Perch. The trimmed mean values, in particular, further supported the robustness of this conclusion by showing that outliers do not skew the results. When compared to other times of the day, the Pav is typically the noisiest in the evenings, with a noticeable increase in noise levels.

Yes, the conclusions in Part V and Part I are the same. As expected, the Buttery was found to be the noisiest location due to its higher and more frequent noise peaks compared to the Pav and the Perch. Although an evening noise peak was observed at the Pav, it was not as loud as the Buttery, which typically has louder noise levels during lunch times. Therefore, this finding does not change the overall conclusion that the Buttery is generally noisier.

Appendix

Code for Matlab Plots

```
close all
clear
warning ('off','all');

% getting data from tables
BMCP = readtable('Buttery morning combinations Plot.csv');
BACP = readtable('Buttery afternoon combinations Plot.csv');
BECP = readtable('Buttery evening combinations Plot.csv');
PaMCP = readtable('Pav morning combinations Plot.csv');
PaACP = readtable('Pav afternoon combinations Plot.csv');
PaECP = readtable('Pav evening combinations Plot.csv');
PeECP = readtable('Perch evening combinations Plot.csv');
PeMCP = readtable('Perch morning combinations Plot.csv');
PeACP = readtable('perch afternoon combinations Plot.csv');


% turning tables into arrays for ease of use
BMCParray = table2array(BMCP);
BACParray = table2array(BACP);
```

```
BECParray = table2array(BECP);
PaMCParray = table2array(PaMCP);
PaACParray = table2array(PaACP);
PaECParray = table2array(PaECP);
PeECParray = table2array(PeECP);
PeMCParray = table2array(PeMCP);
PeACParray = table2array(PeACP);

% Contriving x values from available data, all data same length
x = 1:size(BMCParray, 1);

% Getting y values from data
BMCParrayY = BMCParray(:, 1);
BACParrayY = BACParray(:, 1);
BECParrayY = BECParray(:, 1);
PaMCParrayY = PaMCParray(:, 1);
PaACParrayY = PaACParray(:, 1);
PaECParrayY = PaECParray(:, 1);
PeECParrayY = PeECParray(:, 1);
PeMCParrayY = PeMCParray(:, 1);
PeACParrayY = PeACParray(:, 1);

% plotting
% Buttery
figure('Name', 'Buttery');
subplot (3,1,1);
plot(x, BMCParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Buttery Morning');
grid on;

subplot (3,1,2);
plot(x, BACParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Buttery Afternoon');
grid on;

subplot (3,1,3);
plot(x, BECParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Buttery Evening');
grid on;

% Pav

figure('Name', 'Pav');
subplot (3,1,1);
plot(x, PaMCParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Pav Morning');
grid on;
```

```
subplot (3,1,2);
plot(x, PaACParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Pav Afternoon');
grid on;

subplot (3,1,3);
plot(x, PaECParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Pav Evening');
grid on;

% Perch

figure('Name', 'Perch');
subplot (3,1,1);
plot(x, PeMCParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Perch Morning');
grid on;

subplot (3,1,2);
plot(x, PeACParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Perch Afternoon');
grid on;

subplot (3,1,3);
plot(x, PeECParrayY, 'r--', 'Linewidth', 2);
ylabel('Noise Levels');
title('Perch Evening');
grid on;
```

TEAM CONTRIBUTIONS:

Each member of the group contributed to the project, from recording to data analysis.

We divided the matlab analysis task as follows:

- Plot the datasets - Leo
- Histograms Analysis- Victor
- Central tendencies Calculation - Sadbh

- 5% Trimmed Mean - Lorraine

  20% Trimmed Mean - Suhani

- Conclusion- Lorraine and Suhani