OTOT Task C

🗺️ **Visualization with Python/R/Gephi/Processing/PowerBI**

📜 **Max marks – 3 marks**

📒 **Task:** Create a visualization (at least three different charts) on your choice of dataset. Use any of the following tools: Python/R/Gephi/Processing/PowerBI. If you wish to use any other visualization tool – seek approval by writing to teaching team/clarification channel. Avoid Excel.

StuName Deng Xinrui **StuID** A0274719R **Email**
E1124465@u.nus.edu

I Dataset

1. **Data Origin & Profile**

   🔗 **Link:** https://www.kaggle.com/competitions/birdsong-recognition/data?select=train.csv

   The dataset titled ***Cornell Birdcall Identification*** is hosted on Kaggle and in 2021.

   The dataset is <u>a comprehensive collection of audio recordings and corresponding metadata</u> for the purpose of training and evaluating machine learning models in bird species classification. The dataset, made available through Kaggle, was compiled by experts in the field of ornithology and audio processing.

   The dataset primarily consists of the `train.csv` file, which holds information about each audio recording, including the unique identifier for each recording, the location where it was captured, the timestamp indicating the exact moment of recording, the audio file identifier, and the primary label corresponding to the bird species singing.

   This structured data provides a solid foundation for training machine learning models to recognize and classify birdsongs.

2. **My Point of Interest**

   Compared to directly reflecting certain statistical information, this dataset, although also encompassing a wealth of content, is primarily focused on serving audio purposes. By integrating audio recordings with sample data and leveraging machine learning or deep learning methodologies, this dataset is capable of yielding more specialized and meaningful conclusions regarding computational bioacoustics and avian behavior.

Therefore, my interest in the visualization of this data centers on how to effectively process and extract useful information that can act as a precursor to machine learning or deep learning tasks, ultimately enhancing the serviceability for artificial intelligence missions.

This also aligns with the intrinsic purpose of the dataset, which is to serve as a resource for bird species identification competitions, aiding the algorithmic models of various teams on Kaggle.

3. **Data characteristics**

There are 32 columns of attributes involved in this dataset, and a considerable part of them focus on the audio format and record. The actual analysis does not need to use so many attributes. The following is a selection of the more important attributes to describe.

Attribute Name	Type	Description
ebird_code	Sting	a code for the bird species
author	Sting	the user who provided the recording
location	Sting	where the recording was taken. Some bird species may have local call 'dialects' (can use to seek geographic diversity)
date	String	while some bird calls can be made year round, such as an alarm call, some are restricted to a specific season(can use to seek temporal diversity)
country	Sting	the country of the restaurant
species	Sting	the real name of species of the bird
bird_seen	String	whether can see the bird or not when recording them
sampling_rate	Integer	a term to describe audio file
elevation	Integer	the elevation(m) at which the bird song audio was recorded
length	Integer	the length(s) of the bird's sound
time	String	the specific time when the bird song audio is recorded (24 hours)
description	String	note information for bird song audio is not available for every sample
type	String	the type identification of bird song is relatively rich and related to the behavior and habits of birds
duration	Integer	The duration(s) of a bird song audio sample

*detailed information about the bird codes by appending the code

to `https://ebird.org/species/`

*such as `https://ebird.org/species/amecro` for the *American Crow*

II Purpose of Visualization

1. **Data Details**

The audio recordings themselves, although not directly included in the CSV file, are presumably available separately and can be linked to the corresponding rows in the train.csv using the audio file identifiers. These recordings capture the natural sounds of birds in their natural habitats, offering a realistic and challenging dataset for audio classification tasks.

The dataset's authors are ornithologists and audio processing experts who have carefully curated and annotated the recordings to ensure accurate and reliable labels. The recordings were likely collected over a period of time, spanning different seasons and locations, to capture a diverse range of bird species and singing behaviors.

2. **My Queries**

- How does the duration of bird calls vary with different speeds?

- What is the distribution pattern of bird call durations in the dataset?

- Which countries contribute the most bird call samples in the dataset, and in what proportions?

- Are there any notable differences in the frequency of bird call durations between short and long calls?

- What trends or patterns can be identify in the collection of bird call samples across different countries?

- Is it possible to visualize audio samples of bird calls? What is a good visualization?
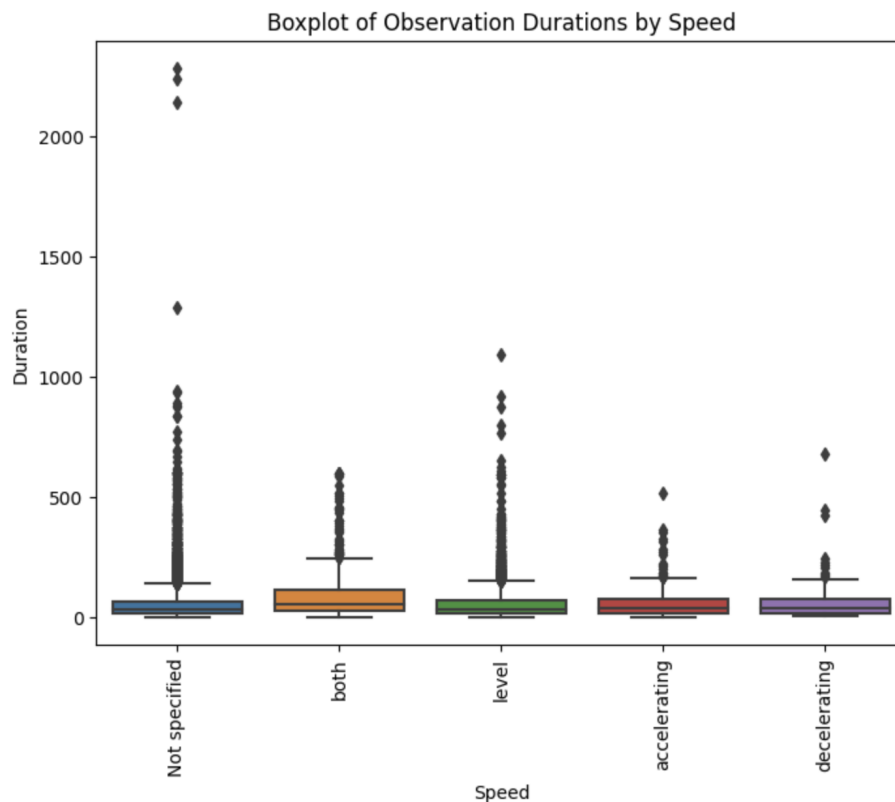
III Visualization

🔗 **Link of Code & Plots in Kaggle** 💚

https://www.kaggle.com/ddsereindd/visualization-of-cbi

🔗 **Link of Code & Plots in Github** 💚

https://github.com/aurora-moon-cloud/VSFG-on-CBI.git

1. **Data Views & Data Encodings**

Box Plot

representing the distribution of observation durations at different speeds.

Boxplot of Observation Durations by Speed

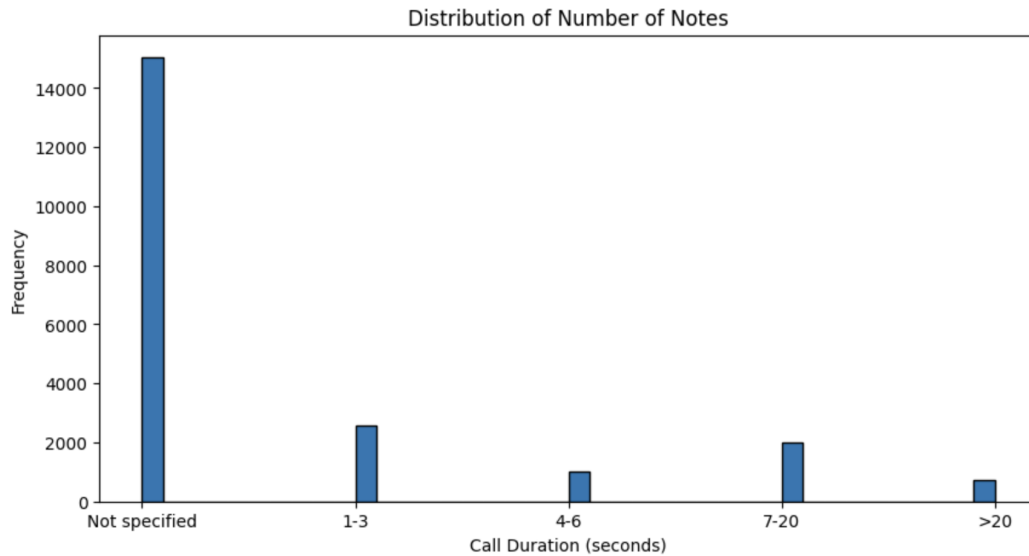The data encodings present in this plot are as follows:

- **Horizontal Axis**: The horizontal axis encodes speed, with different categories represented by different boxes in the plot.

- **Vertical Axis**: The vertical axis encodes observation duration, with longer durations represented by higher positions on the axis.

- **Legend**: The legend provides an additional encoding, indicating which subplot corresponds to which scenario or dataset.

The center line inside each box represents the median duration, the top and bottom edges of the box represent the upper and lower quartiles (25% and 75% percentiles), respectively, and the whiskers extend to the most extreme data points within 1.5 times

the interquartile range. Any data points beyond this range are represented as individual dots.

Histogram

a graphical representation that illustrates the frequency distribution of numerical data.


Distribution of Number of Notes

The data encodings included in this histogram are as follows:

- **Bar Height**: The height of each bar in the histogram represents the frequency or count of calls with a specific duration. The taller the bar, the more frequent those calls were with that particular duration.

- **Bar Width**: The width of each bar represents the range of call durations it covers. In this case, each bar represents a specific duration range, such as 0-5 seconds, 5-10 seconds, etc.

- **X-axis**: The horizontal axis (X-axis) represents the call durations. It is typically labeled with increasing values indicating the duration in seconds or minutes.

- **Y-axis**: The vertical axis (Y-axis) represents the frequency or count of calls. It is labeled with numerical values indicating the number of calls that fall within each duration range.

Pie Chart

representing the distribution of bird call samples collected from different countries.
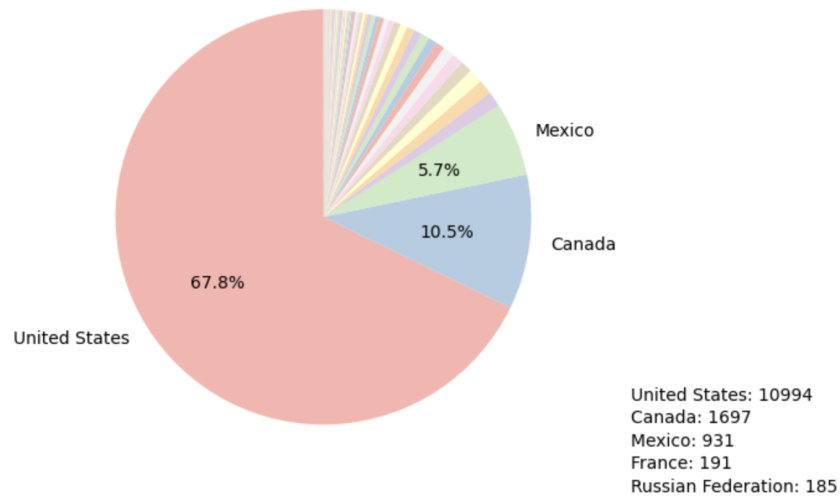
The data encodings included in this pie chart are:

- **Country Representation**: Each slice of the pie represents a specific country, where the bird call samples were collected. The size of each slice is proportional to the number of samples collected from that country.

- **Percentage Labels**: The chart includes percentage labels next to each slice, indicating the proportion of samples contributed by each country. These labels provide a quantitative understanding of the distribution.

- **Color**: Each country is represented by a different color, which helps distinguish between the various slices of the pie. The color coding makes the chart visually appealing and enhances the readability of the data.

Mel spectrogram

a visualization tool used to represent the acoustic features of sound signals, specifically in this case, the calls of a White-crowned Sparrow.

It provides a frequency-time representation of the audio signal, with the Mel scale used to map frequencies onto a more perceptually relevant scale for humans.

It approximately represents the sound wave patterns and frequency content of the bird's call over time. It captures both the temporal structure of the call, indicating when different frequencies occur, and the spectral structure, indicating the relative strengths of different frequencies at any given moment.
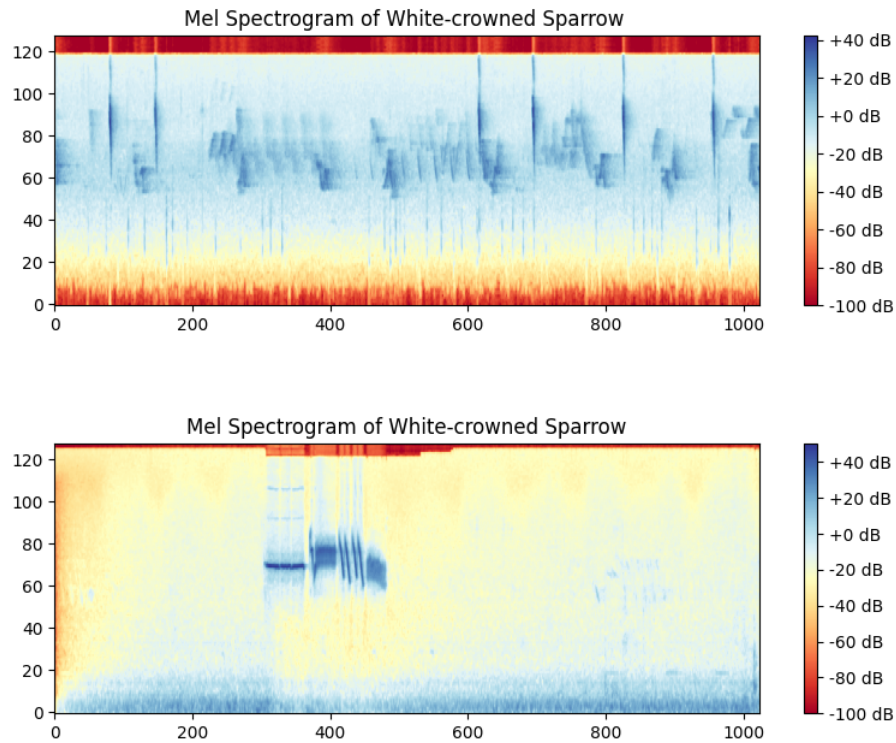
The data encodings present in this Mel spectrogram are:

1. **Frequency**: The vertical axis represents frequency, with the Mel scale used to transform linear frequency into a more psychologically relevant scale. This encoding allows us to visualize the frequency content of the bird's call, indicating which frequencies are dominant at different times.

2. **Time**: The horizontal axis represents time, with the spectrogram showing how the frequency content changes over time. This encoding enables us to understand the temporal structure of the bird's call, identifying any patterns or modulations that occur over time.

3. **Amplitude**: The color or intensity of each point in the spectrogram represents the amplitude or loudness of the sound at a particular frequency and time. This encoding

provides information about the relative strengths of different frequencies at any given moment, indicating which frequencies are louder or softer.

2. **Findings & Answers**

Upon observing the box plot, we can conclude that observation durations vary significantly depending on the speed. Some speeds may have a narrower distribution of durations, indicating more consistency in observations, while others may have a wider distribution, indicating more variation. Additionally, certain speeds may have systematically longer or shorter observation durations than others.

Upon observing this histogram, it appears that the majority of calls are relatively short, with most of the bars concentrated towards the left side of the graph. This suggests that most songs of bird last for a few seconds or minutes. However, there are also a few songs that last significantly longer, as indicated by the taller bars on the right side of the graph. These longer calls, although less frequent, still deserve attention as they might represent exceptional or outlier situations.

Upon observing the pie chart, it becomes evident that the majority of bird call samples in the dataset were collected from the United States, followed by France. This suggests that these countries may have a more diverse range of bird species or a higher concentration of bird habitats, making them prime locations for bird call sample collection.

The relatively smaller slices representing Mexico, Canada, and Russia suggest that fewer samples were collected from these countries. This could be due to limited bird habitats, fewer bird species, or less extensive sampling efforts in these regions.

Upon comparing the spectrograms of the calls of White-crowned Sparrows, significant differences are evident. These disparities can be attributed to various factors such as environmental variations during sampling, behavioral states of the birds, individual differences, recording equipment, and recording conditions. These factors potentially influence the frequency, amplitude, and structure of the sounds captured, resulting in distinct spectral patterns.

The above information lays a foundation for the further study and research of artificial intelligence algorithm model.