Today let us introduce an application of Fourier Transform – Rainforest connection species sound detection.
First, let me introduce some essential background of this problem and have a brief view of the dataset.

The presence of rainforest species is a good indicator of the impact of climate change and habitat loss. Hearing these species is more straightforward than seeing them, so it's important to use acoustic technologies that can work on a global scale. Additionally, real-time information, such as that provided through machine learning techniques, could enable **early-stage detection of human impacts** on the environment. This result could drive more effective conservation management decisions.

And while **automatic acoustic identification via deep learning** has been successful, models require a large number of training samples per species. This limits applicability to ***rarer species***, which are central to conservation efforts. Thus, methods to *automate high-accuracy species detection in noisy soundscapes with limited training data* are the solution.

Now let us briefly check the dataset.

There are 7 columns, including recording_id, the record file name, species_id, the identifier for different species, the songtype_id, the identifier for song type, and we use t

to describe the range of start and end second of the annotated signal, and use f to describe the range of frequency of the signal.

Next, let us do the target count distribution analysis. Here is the training set. We can see that number of records of each species is almost the same, and for most of the species with one song type, only 17 and 23 species have two types of song.

Next, we separately analyze some data properties based on species and song type. Clip duration is used to describe each species' length and interval of their appearance in audio. As we can see, most species have a value near 2. And here is their amount and percentile. And when calculated based on song type, we can see that the number of type 1 has a relatively small value compared with type 4.

Then, we analyze the frequency range of different species and song types. This variable is used to describe species' sound frequency interval. It is helpful because different species have different frequencies, and we can be based on such properties to identify species. Most species' length of the frequency range between 2000 and 4000. And we can also conclude that type 4 has a broader range than type 1.

In this project, we first analyze the characteristics of the data and audio, use the short-time Fourier transform to map the data from the time domain representation to the frequency domain representation, and then extract the spectral centroid and spectral roll-off features of the sounds of different species.

Finally, the audio signal features are extracted using the Mel coefficient, and other noises are removed. We show species-specific waveforms and frequencies in mixed audio via visualizations such as Mel-spectrograms. We will identify the sounds of different species in an efficient way and provide help for the subsequent analysis of the sounds of different species using deep learning and other methods, triggers more effective species conservation decisions.