

"Cheep Shots": Investigating Bird Populations with Audio Classification

COMP 4449 Capstone Final Project

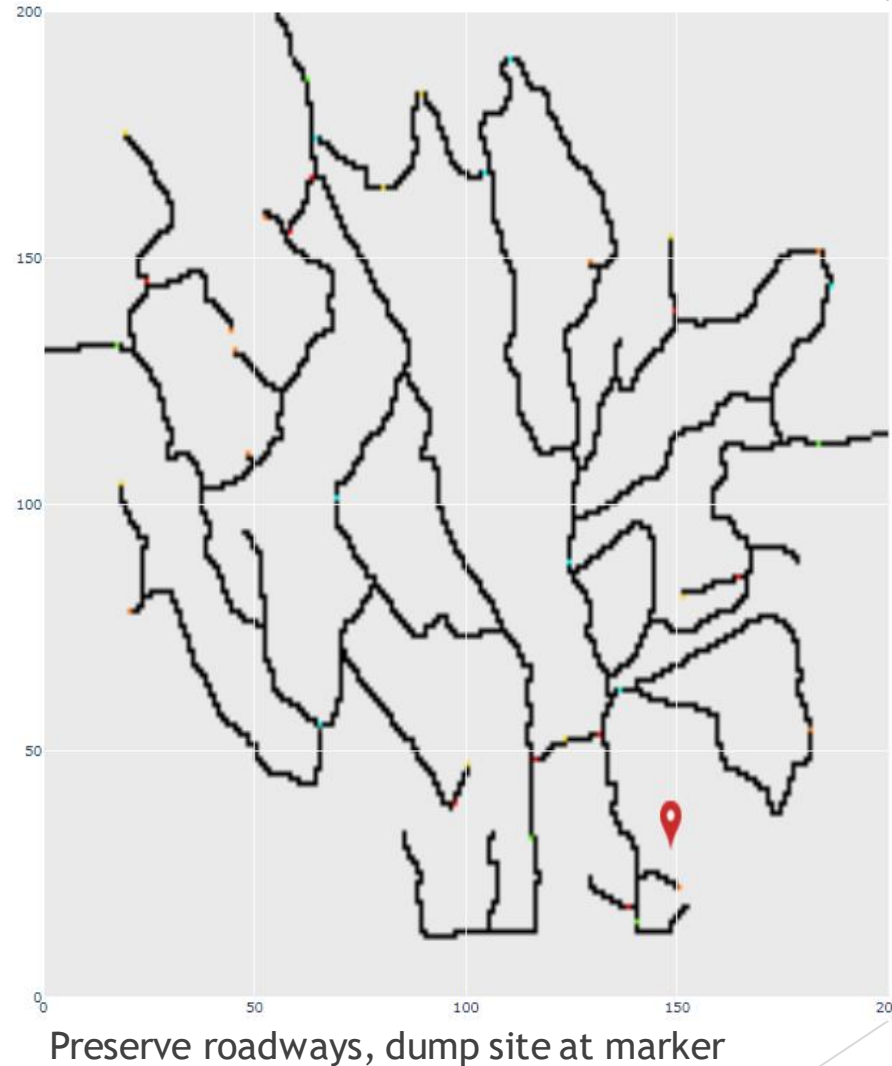
Luke Sonnanburg

Contents

- ▶ Background
- ▶ Problem Statement
- ▶ Data
- ▶ Methodology
 - ▶ Pre-processing
 - ▶ Data Exploration
 - ▶ Feature Engineering
 - ▶ Analysis
- ▶ Results
- ▶ Conclusion

Background

- ▶ Wildlife Preserve monitors bird populations with audio recordings
- ▶ Kasios Office Furniture manufacturer dumps harmful chemicals in Preserve
- ▶ Number of rose-crested blue pipit recordings plummets
- ▶ Kasios accused of harming pipits
- ▶ Kasios produces recordings of pipits to prove birds are fine



Problem Statement

Can we use the available information to appraise the well-being of local bird populations and the manufacturer's culpability?

And the core of answering that question:

Are the Kasios rose-crested blue pipit recordings authentic?

Data

The Preserve provided:

- ▶ Roughly 2,000 audio recordings of birds
 - ▶ 15 different identified species
 - ▶ Most recordings either songs or calls
- ▶ Metadata about recordings
 - ▶ Location, quality (A-D), date and time
- ▶ Map of preserve roadways

Kasios provided:

- ▶ 15 recordings they claim are of rose-crested blue pipits
- ▶ Associated metadata for recording location

Pre-processing/Data Cleaning

- ▶ Ensuring all recordings are in the correct geographic area
- ▶ Removed recordings marked as low-quality (below "B")
- ▶ Removed recordings that weren't "call" or "song"
- ▶ Removed recordings with invalid dates

Data Exploration

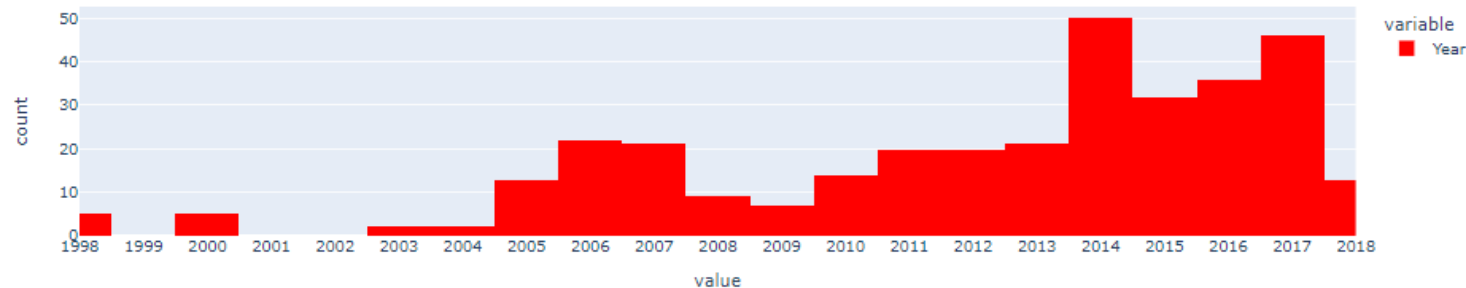
- ▶ 181 of nearly 2000 recordings are target species



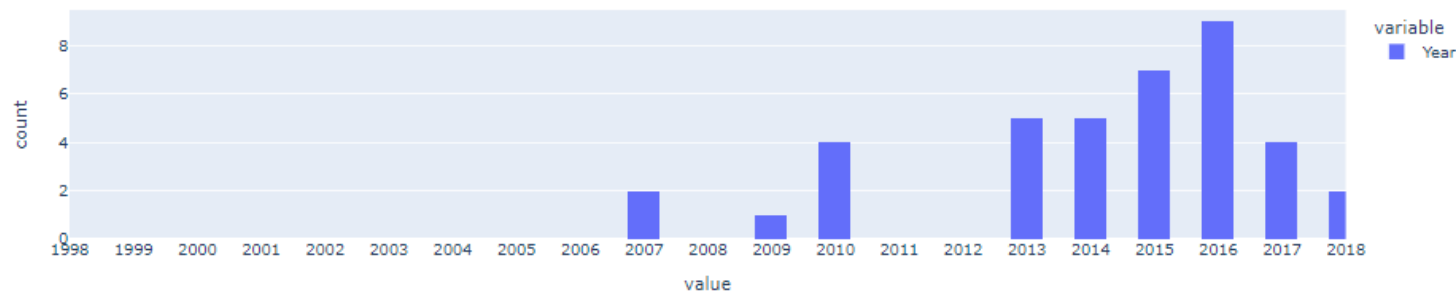
Data Exploration

- ▶ 2018 data only goes through March; comparing January-March data for all years
- ▶ Massive dip in both graphs for 2018, downward trend starts in 2016 for pipits
- ▶ Problem did not state when dumping occurred

Count of all bird recordings in early months by year

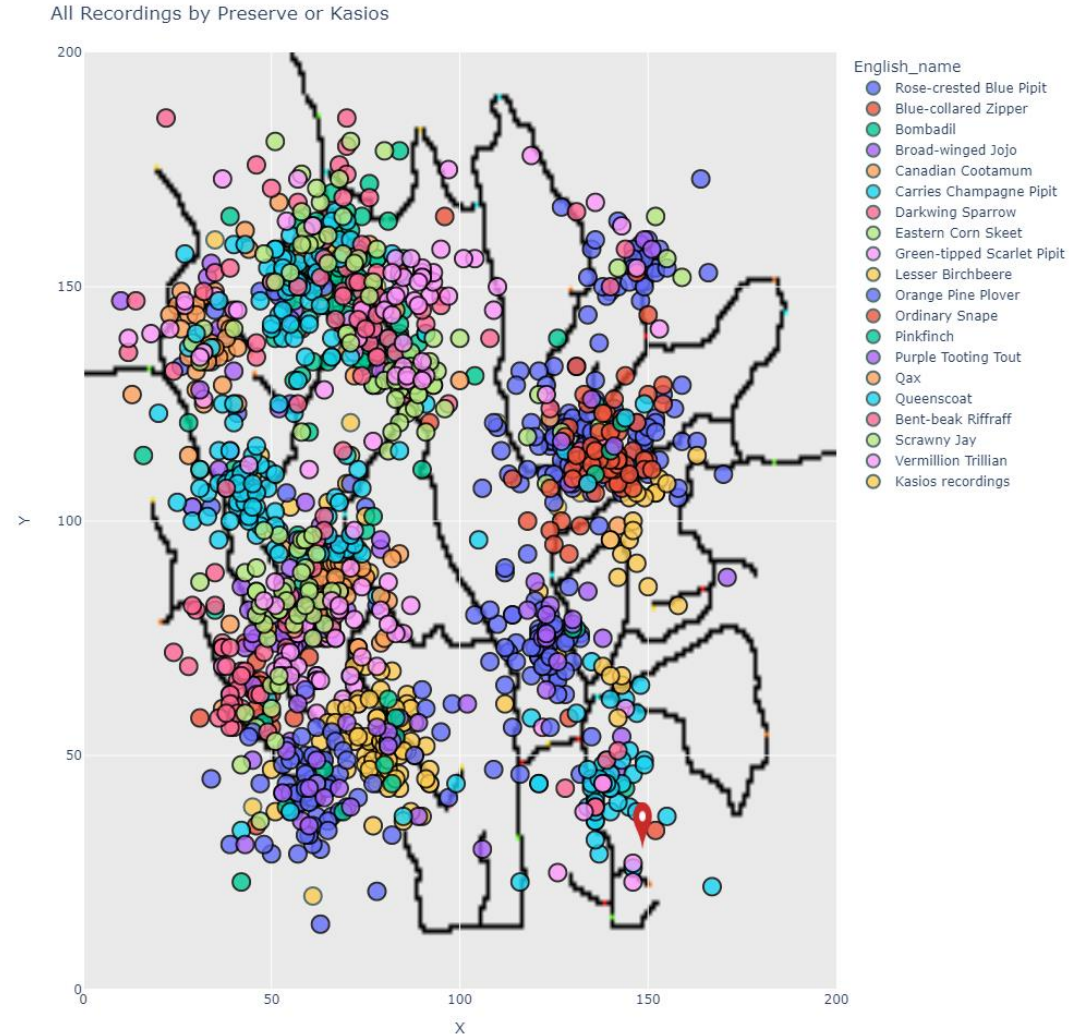


Count of Rose-crested Blue Pipits in early months by year



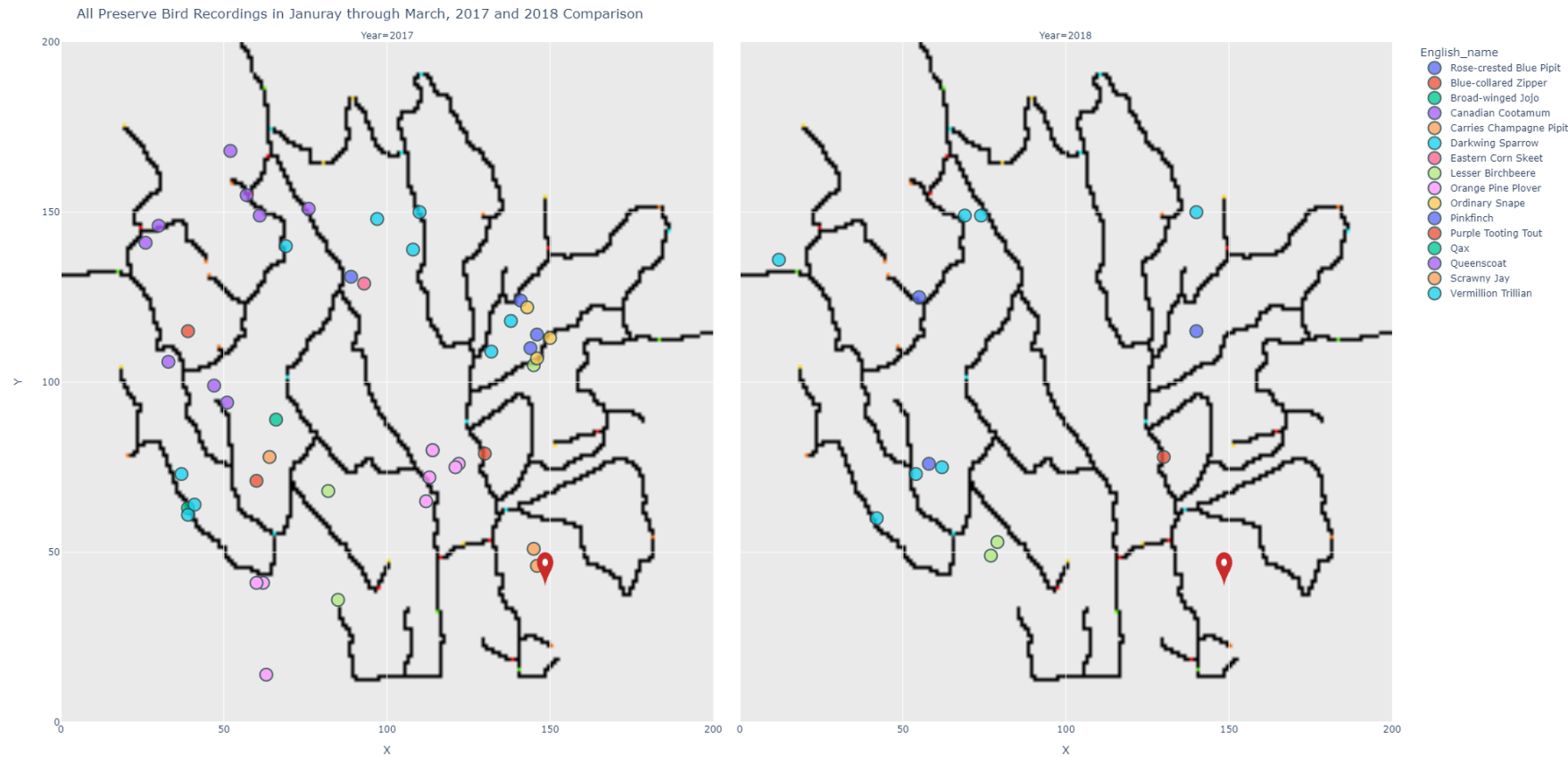
Data Exploration: Geo Data

- All recordings plotted
- Most species concentrate in bounded spaces



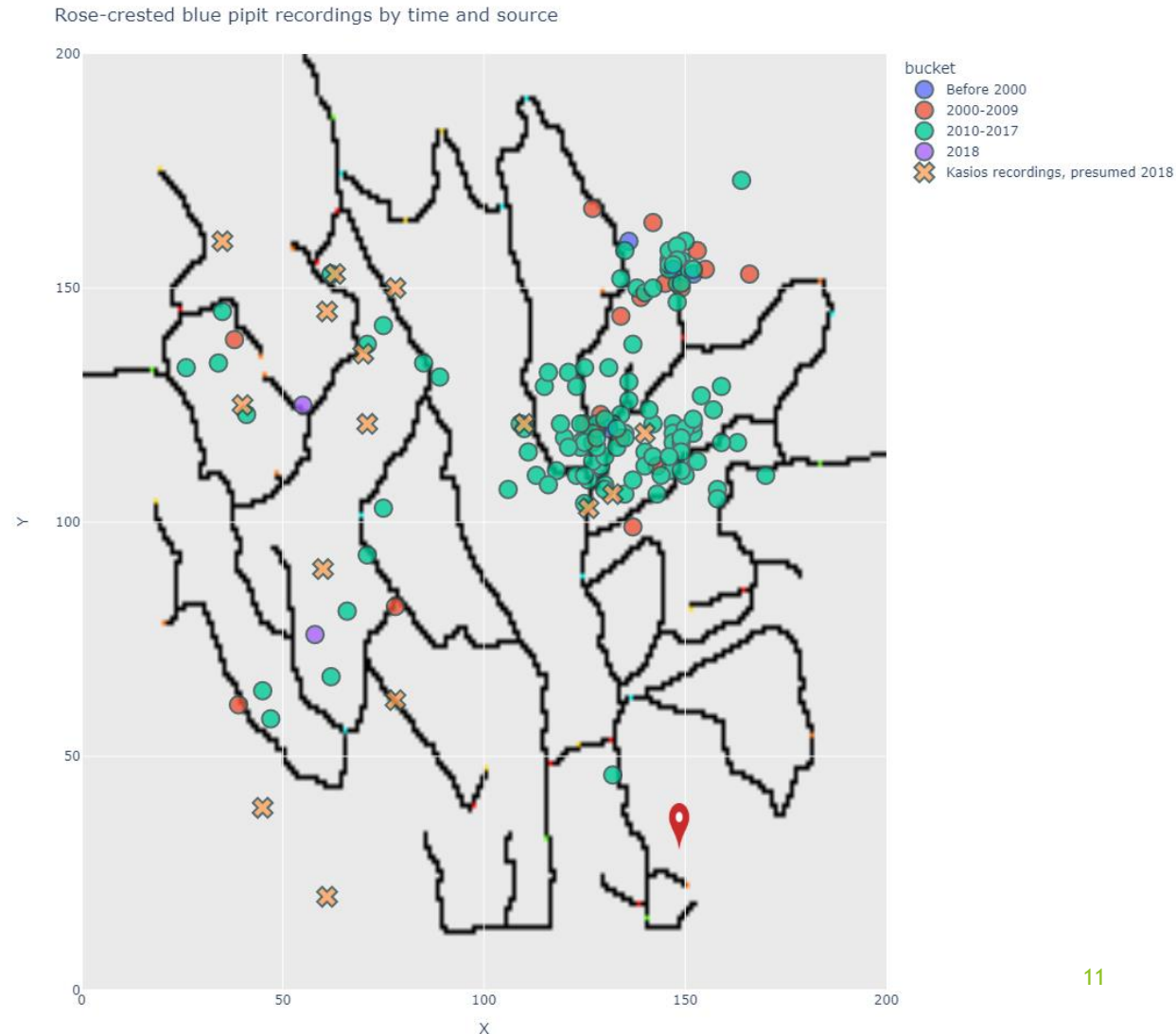
Data Exploration: Geo Data

- Decrease in early month recordings from 2017-2018 from
- No recordings very near dump site



Data Exploration: Geo Data

- ▶ Recent pipit recordings mostly concentrated in northeast
- ▶ The two 2018 recordings are scattered further from dump site
- ▶ Kasios recordings seem to be randomly distributed, with only four near typical pipit habitat
- ▶ Simple classifier applied to this data later



Understanding Audio Data

- ▶ Computer reads in "amplitude" or extent of air particles oscillating due to sound
 - ▶ Python's librosa library loads audio data and extracts features
- ▶ Loaded with a "sample rate"; how many amplitude samples per second
 - ▶ Sample rate 24,000 supposed to be good for bird frequencies
- ▶ Fourier transform yields magnitude of frequencies in signal
- ▶ Spectrogram graphs show magnitude of different frequencies over time
- ▶ Classifying graphs becomes an image classification task

Handling Audio Data: Feature Selection

Librosa feature extraction gets more specific data from recording

Features used here:

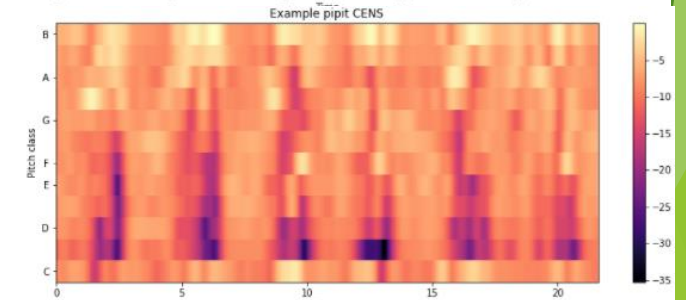
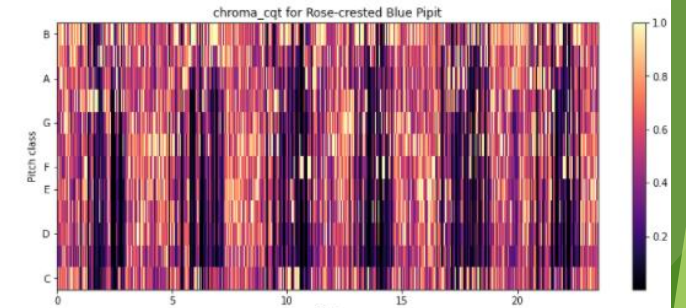
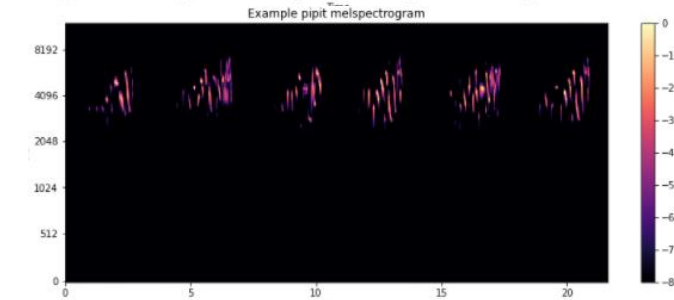
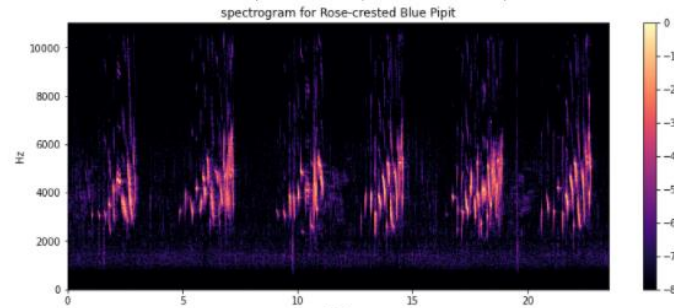
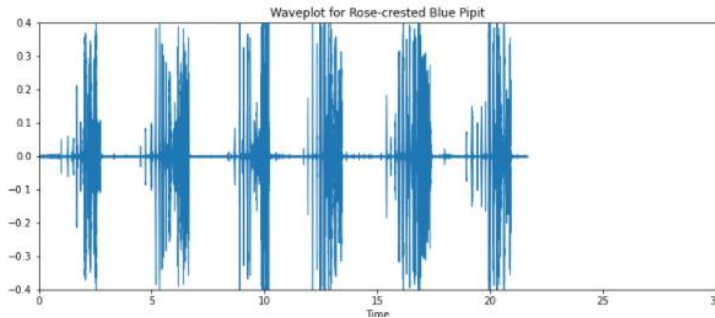
- ▶ Root mean square of amplitude: Average "loudness" of sample
- ▶ Chroma energy normalized/cens: Pitch over time; captures melodic progression
- ▶ Melspectrogram: Pitch over time scaled to human hearing
- ▶ Spectral-roll off frequency: Sort of measures audio skewedness; frequency below which 85% of power lies
 - ▶ Samples with sharp, high frequency noises like cymbals will have high SRF
- ▶ Zero Crossing Rate: Number of times signal passes through zero
 - ▶ Can be thought of as smoothness; noisy backgrounds will cause high ZCR

Visualizing Audio Data

Certain data that's better for computer classification is messier for humans to read.

Used in visualization:

- ▶ Waveplot, shows amplitude over time
- ▶ Spectrogram, shows frequency over time,
- ▶ Constant-Q chromagram shows normalized frequency scaled to pitch over time; melodic progression of birdsong

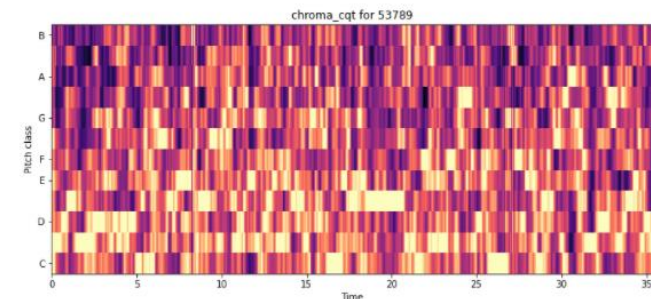
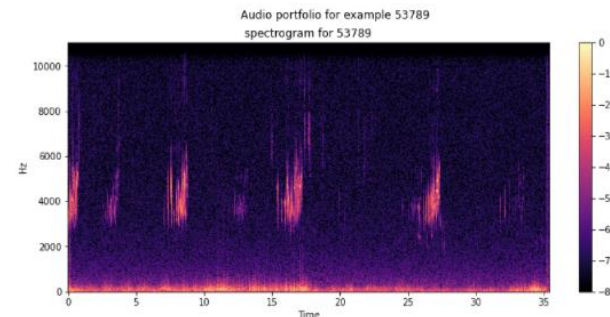
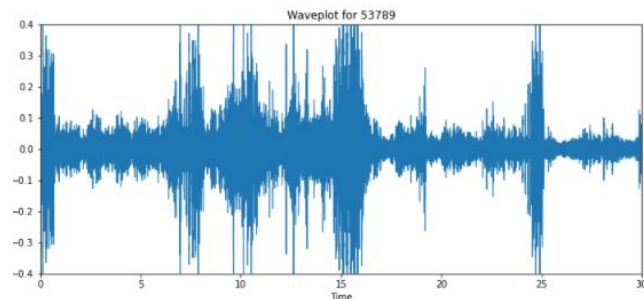
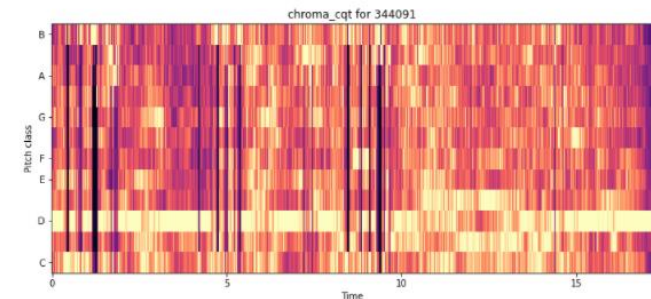
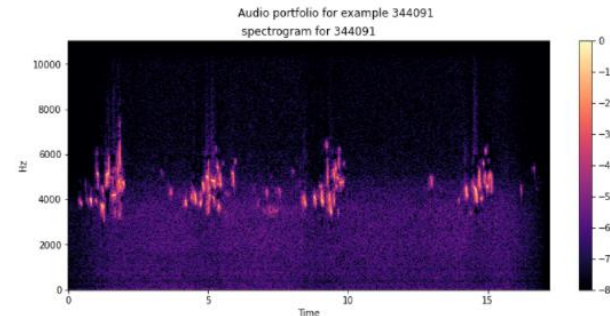
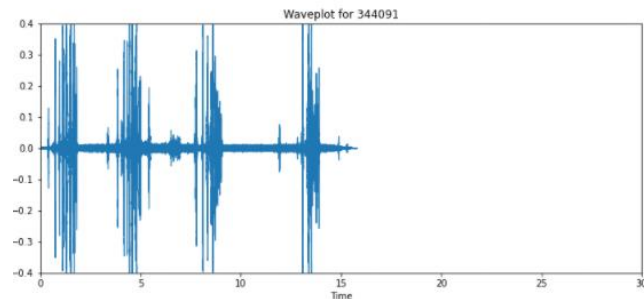
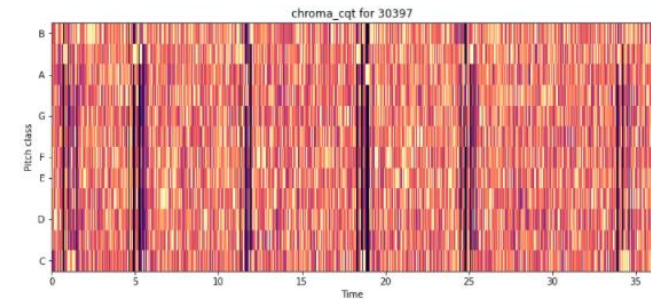
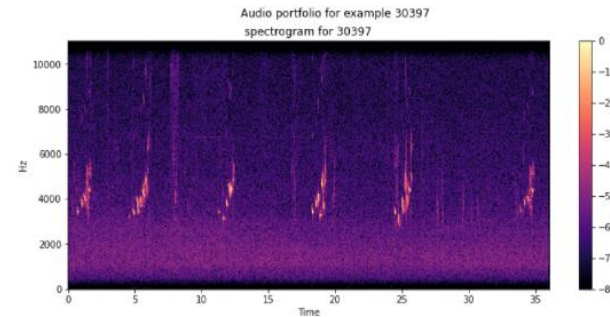
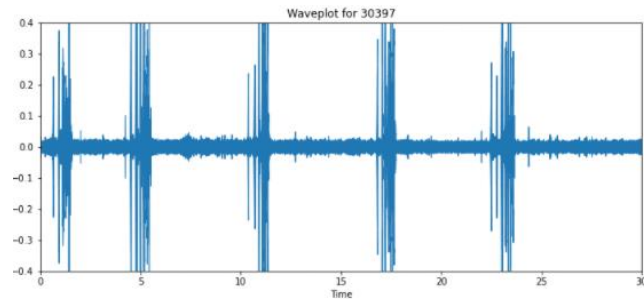


Used in classification but harder for humans to read:

- ▶ Melspectrogram, shows frequency over time scaled to human hearing,
- ▶ CENS, chromagram scaled and normalized into a blurry mess to be more robust against noise pollution and fluctuations

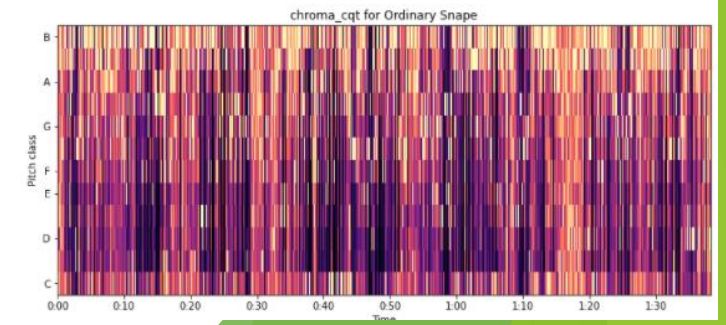
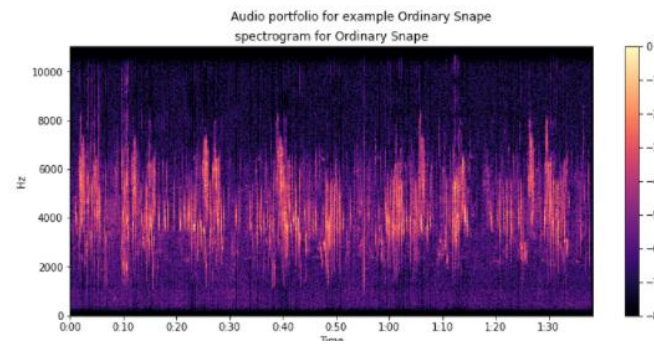
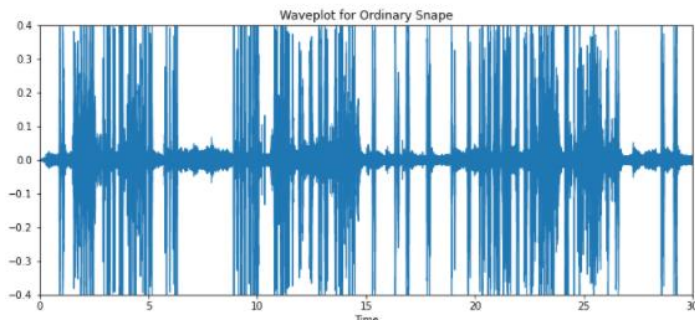
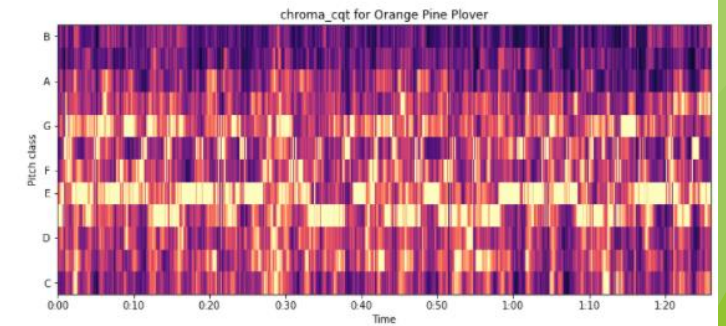
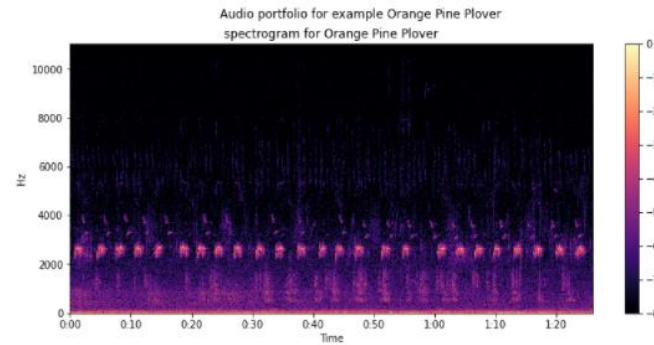
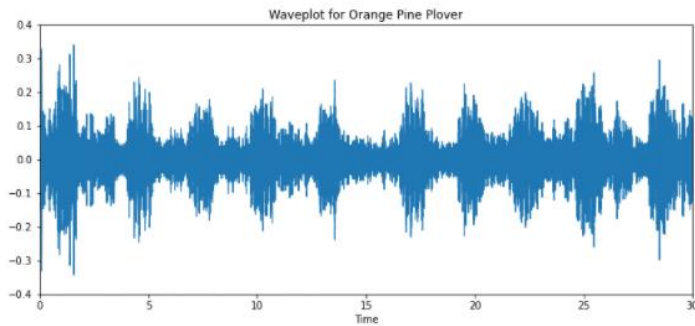
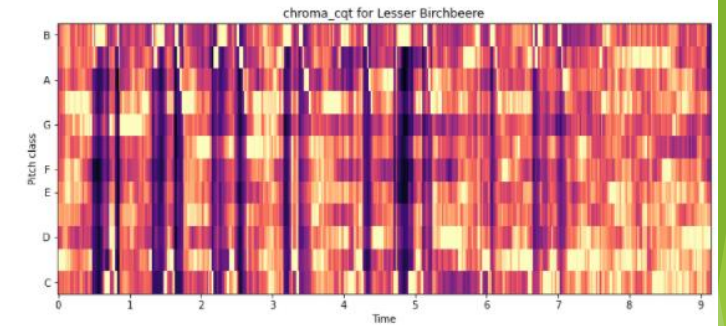
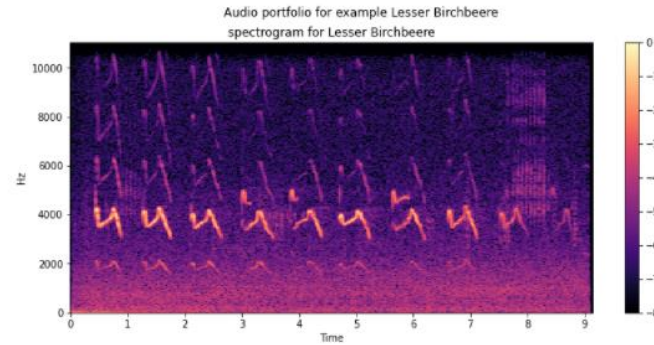
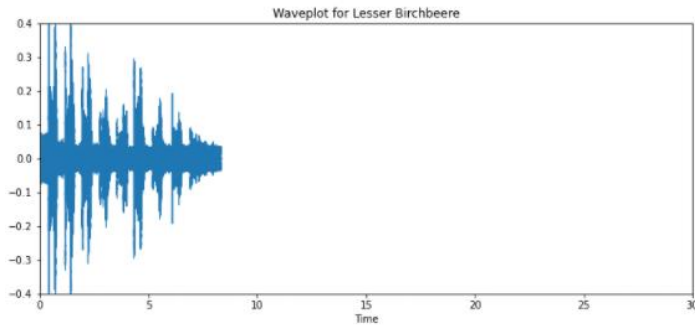
Audio Visualization for Pipits

- ▶ Short, escalating trills that get higher in frequency, about 3-5 seconds apart
- ▶ Vocalizations between 3-6 kHz



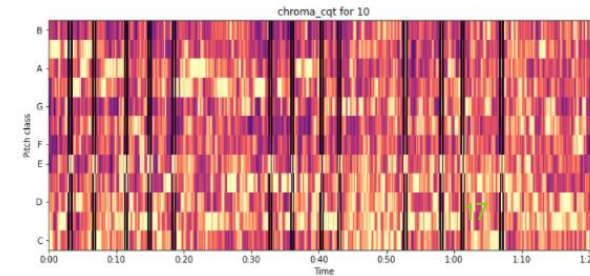
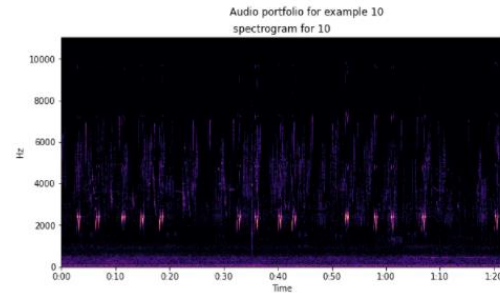
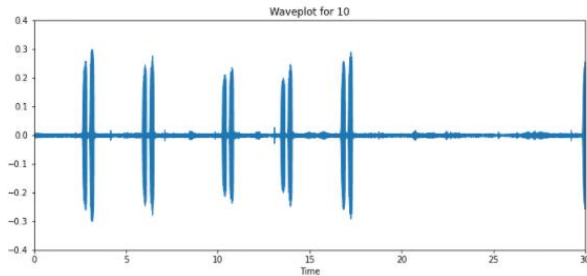
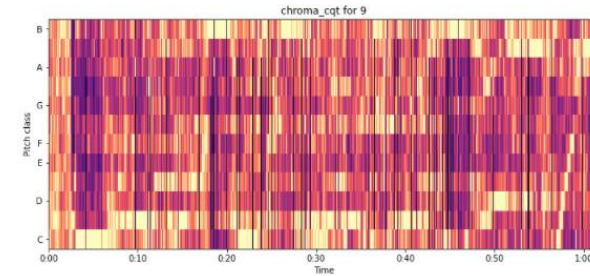
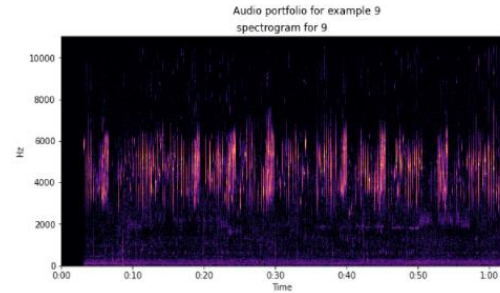
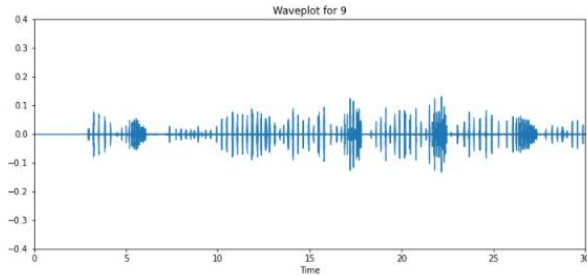
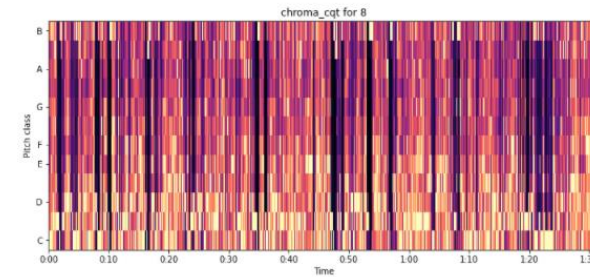
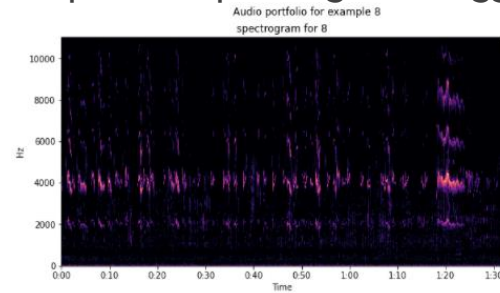
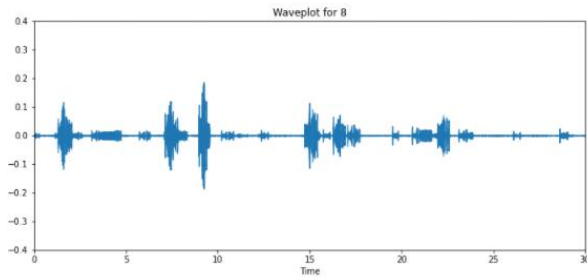
Audio Visualization: Various Species' Songs

No particular consistency in any graphs between species



Audio Visualization for Kasios Recordings

- ▶ No significant consistency
- ▶ Recording 9 somewhat resembles the patterns of rose-crested blue pipit recordings
- ▶ Flat amplitude and abundant dead space in spectrograms suggests unusually clean recordings



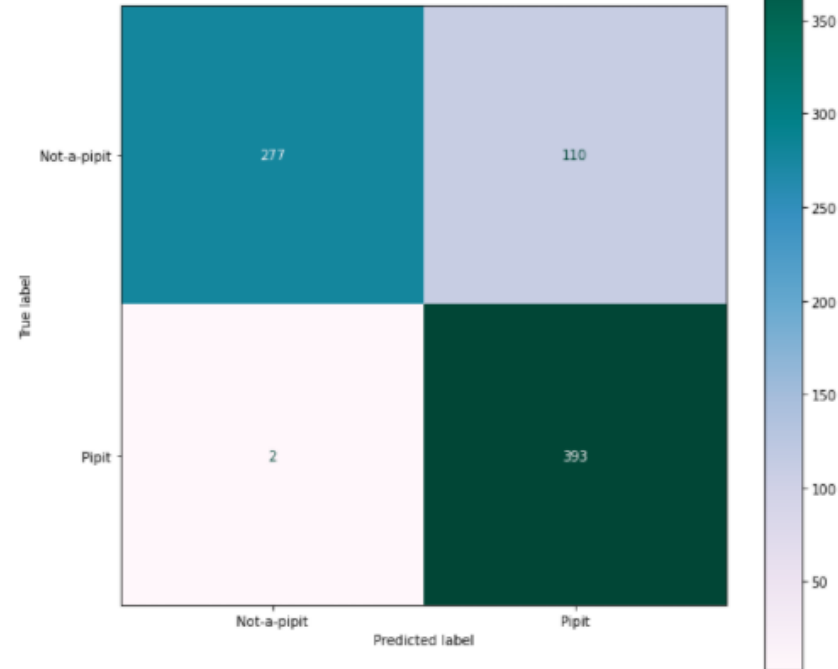
Model Selection

- ▶ Only care about rose-crested blue pipits; binary classifier for classes "pipit" and "not a pipit"
- ▶ Resulted in "pipit" class only being about 10% of data set
 - ▶ SMOTE oversampling used to balance classes
- ▶ Tried fitting several models in cross-validation pipeline
- ▶ Winning model chosen based on highest recall with reasonable accuracy
- ▶ Recall prioritized to give Kasios greatest benefit of the doubt, avoid any risk of false accusations
- ▶ Winning model: K nearest neighbors
- ▶ Similar steps were taken with the very similar geographic data

Model Performance

- Model is fairly reliable with overall 0.86 accuracy
- 0.99 recall for rose-crested blue pipits is almost perfect; should be very generous to Kasios
- Similar steps for the geographic data were taken with 0.91 recall

	precision	recall	f1-score	support
Not-a-pipit	0.99	0.72	0.83	387
Pipit	0.78	0.99	0.88	395
accuracy			0.86	782
macro avg	0.89	0.86	0.85	782
weighted avg	0.89	0.86	0.85	782



Model Results

- ▶ Audio classifier finds that only the 13th recording is a rose-crested blue pipit
- ▶ Geographic classifier finds that only the 15th recording came from a likely pipit habitat

Recording	Audio	Geographic
1	Not a pipit	Unlikely pipit
2	Not a pipit	Unlikely pipit
3	Not a pipit	Unlikely pipit
4	Not a pipit	Unlikely pipit
5	Not a pipit	Unlikely pipit
6	Not a pipit	Unlikely pipit
7	Not a pipit	Unlikely pipit
8	Not a pipit	Unlikely pipit
9	Not a pipit	Unlikely pipit
10	Not a pipit	Unlikely pipit
11	Not a pipit	Unlikely pipit
12	Not a pipit	Unlikely pipit
13	Pipit	Unlikely pipit
14	Not a pipit	Unlikely pipit
15	Not a pipit	Likely pipit

Conclusion

Kasios is attempting to prove its innocence by producing 15 recordings of rose-crested blue pipits, however...

- ▶ Audio classifier determines only one recording to be from a pipit
- ▶ Geographic data makes only four recordings look sourced from likely pipit habitats, classifier says only one is likely pipit habitat
- ▶ Human interpretation of spectrograms and audio playback makes most recordings seem like they came from different species
- ▶ Unusually clean audio casts some doubt on whether these recordings even came from the Preserve

I can't confirm or deny that Kasios's illegal dumping caused a population decline or displacement.

However, Kasios was either dishonest or careless in building their case and cannot be trusted to investigate themselves.