

# Report on Assignment 1

## HashMap Team

### Goal of the assignment:

The goal of this first assignment working on the VAST Mini Challenge 1 is to visualize the map of the Preserve housing the different species of birds and to add to this map the bird locations we have.

### Step By step implementation:

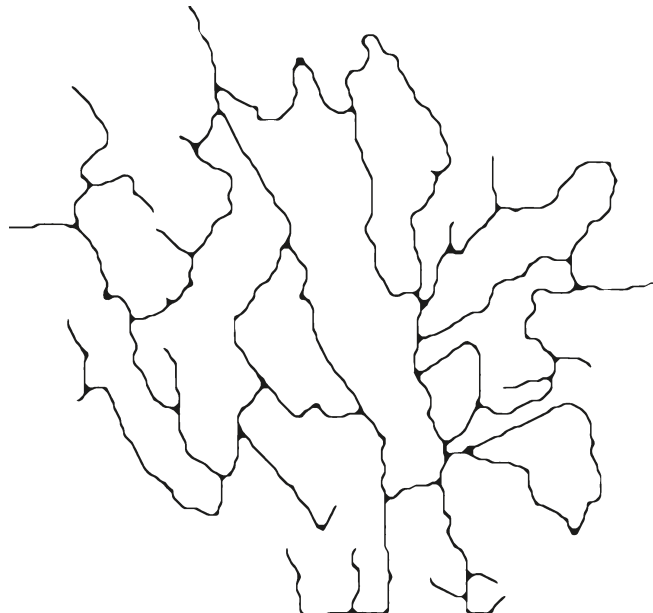
We started the assignment by simply plotting the original map of the Preserve. The result was the following image to the right.

We can notice two distinct problems with this map: First of all, the original map is quite small and trying to resize it using python would leave us with a sketchy image of very low resolution. Second, the original map contains a plan of roads on the site. The edges of the roads are blurry and some pixels are the of the wrong color. The contrast is also off.



*Original map of the Preserve*

To address these two problems, we transformed the original map using a combination of python tools as well as external softwares to have a better rendering of the Preserve's geography.



*Processed map of the Preserve*

Now that we have a usable map, we can move on to plotting the different bird locations on the Map.

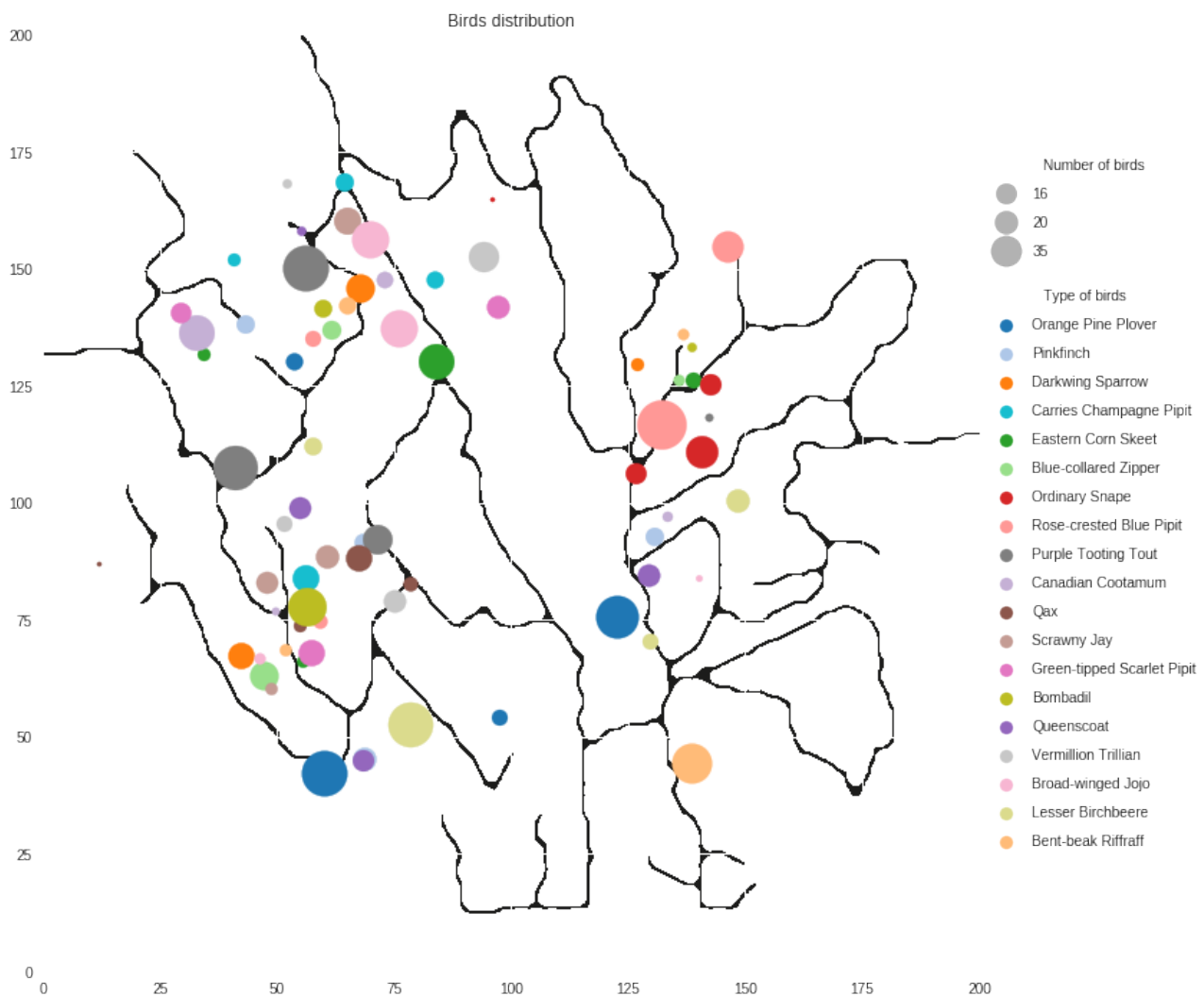
We notice that we are given over 2000 bird locations from around 20 different species in the 'AllBrids.csv' File. So instead of plotting all the birds on one map we decided on the following strategy:

We will use the 2D map as a background for our plots:

- We will represent every bird species with a different color as we are dealing with different categories.
- Instead of plotting every bird coordinates as a dot (or with an 'o' Marker ) on the map, we will plot clusters of birds that we will compute using a clustering algorithm. The size of the ( 'o' Marker ) representing a cluster will be proportionate to the population of birds of that cluster. The coordinates of the cluster will be computed based on the different coordinates.

This way, we avoided having too many dots on the map.

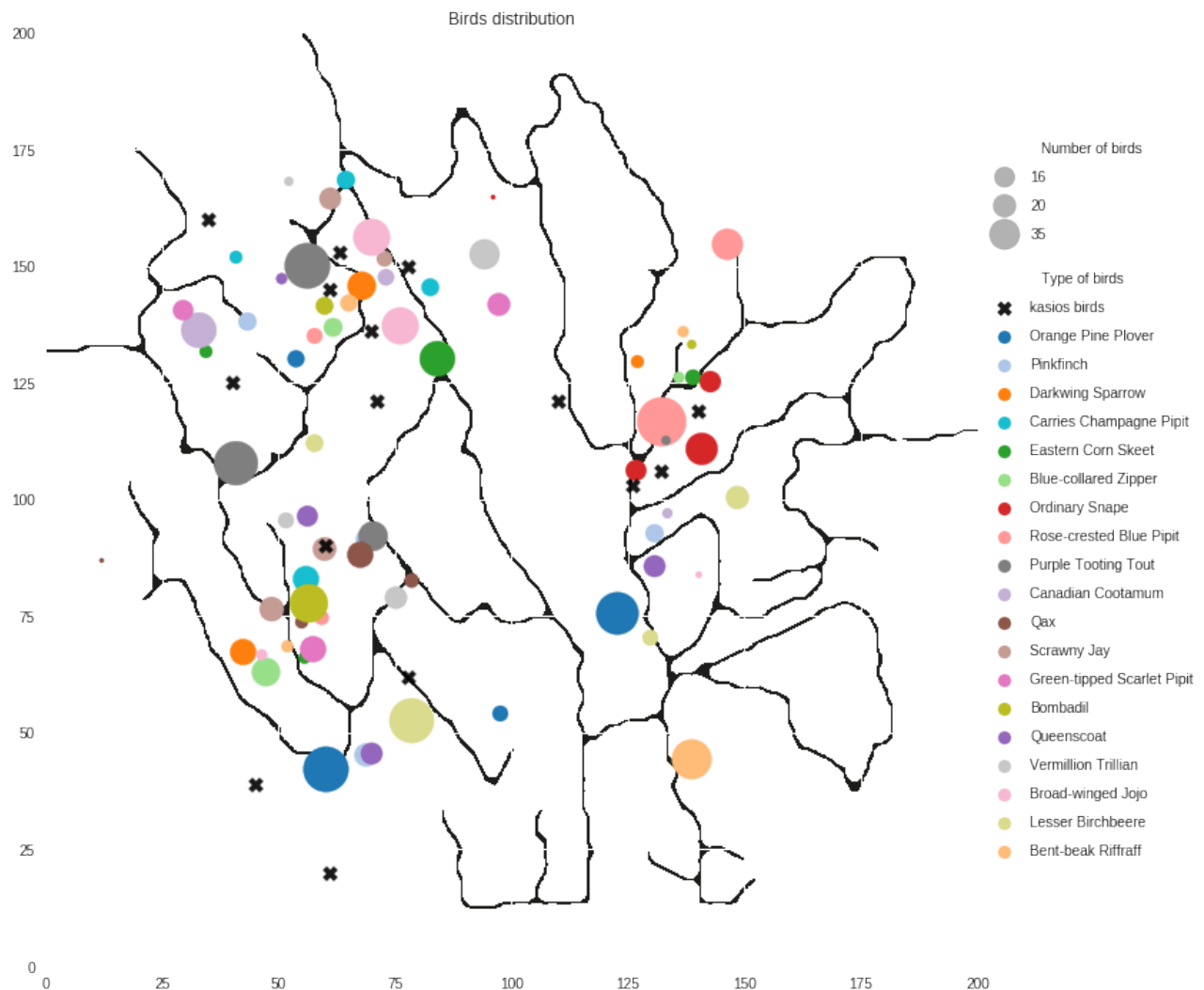
The results of our clustering algorithm and cluster plot is the following:



*Cluster representation of the different bird species ( Rose crested blue pipits in Pink )*

To this representation, we added the 15 location of birds provided by Kasios and supposed to be Rose crested Blue pipits to the map. We marked there locations with a cross to distinguish them from the other known birds.

The final result is the following:



*Cluster representation of the different bird species with the kasios bird locations*

Our conclusion regarding this first assignment are the following:

- Using the clusters allowed us to get a general feel of the distribution of the different species, and more specifically the Rose Crested Blue Pipit which seem to hover around the north est side of the Preserve as indicated by the pink colored clusters.
- The birds from Kasios have varying locations all over the preserve which seem to indicate that not all birds are Rose Crested Blue Pipits, or that in the least, the birds have migrated from their original spot on the land. This puts in doubt the alleged return of the Blue pipits as the effects of the hazardous dump might not have cleared with time.

Our representation have certain downsides:

- Using clusters instead of individual locations generate a loss of precision of the birds locations on the maps: indeed, we are left with only general guidelines as to where the birds have been recorded with no precise location information.

- Displaying each different bird specie in a color is too overwhelming since we can't locate the Blue Pipits easily.

### **Next Steps:**

After analyzing the pros and cons of our initial strategy, we imagine the following steps to our analysis to be as followed:

- Consider using a library that enable adding and removing data point from the maps based on different criteria such as time, coordinates, bird specie, etc ... We might consider the Bokeh library.
- Highlight the Rose crested Blue Pipits locations compared to the rest of the species to investigate the differences in localization between known Blue Pipits and the Kasios test set.
- Investigate the importance of the localization of a bird specie in classifying the Kasios birds, and come up with a first result on whether the provided birds from Kasios are Blue Pipits or not.
- Combine this first approach with the sound file analysis to generate more accurate classification of the bird species.