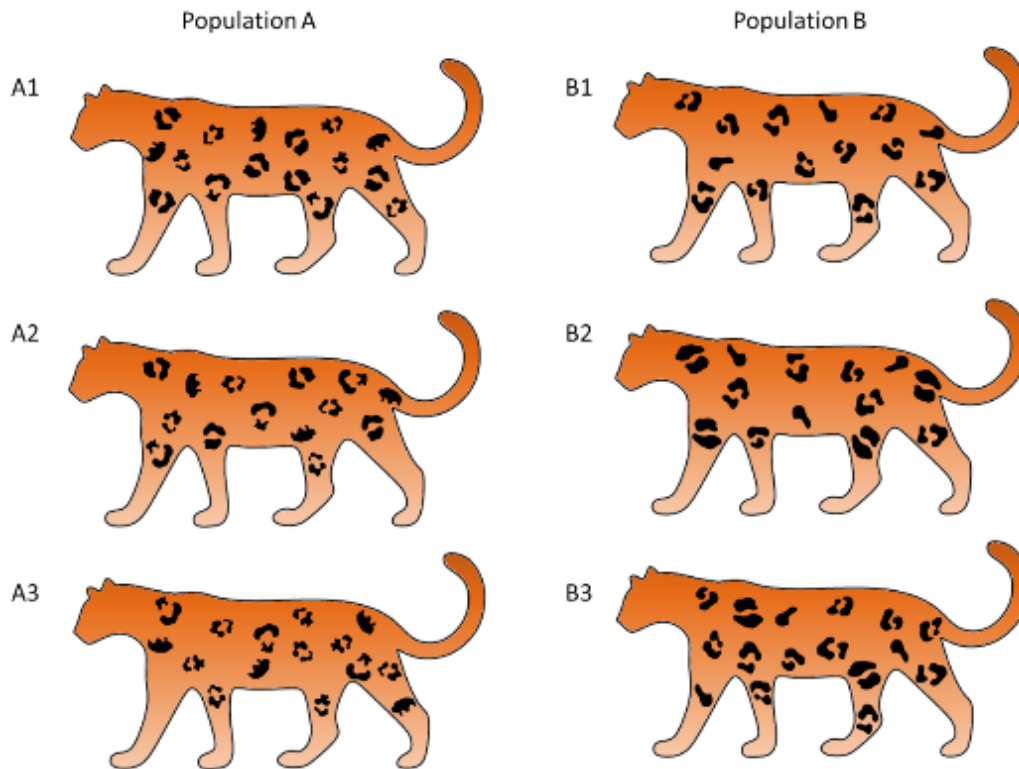


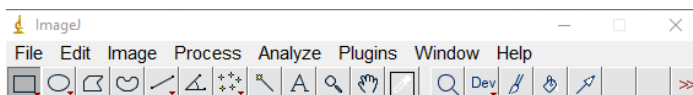
Aim

In this activity, we're going to try to distinguish between the patterns (and only the pattern) on two populations of leopards. Just by looking at the pictures below, can you identify specific differences between the populations or predict which population a specific individual belongs to? This is where pattern analysis software can help us and we're going to see how.

Installing programmes

The first thing we need to do is to install all the programmes we're going to use. ImageJ is the main programme, and PAT-GEOM and Fourier Shape Analysis run inside ImageJ.

- 1) Install ImageJ
 - a. Go to: <https://imagej.nih.gov/ij/download.html>
 - b. Download the version for your OS and follow the instructions
- 2) Install PAT-GEOM
 - a. Go to: <http://ianzwchan.com/my-research/pat-geom/>
 - b. Download "PAT-GEOM.zip"
 - c. Copy and paste the entire contents into the Plugins folder of your ImageJ installation
- 3) Install Fourier Shape Analysis
 - a. Go to: https://imagejdocu.tudor.lu/doku.php?id=plugin:analysis:fourier_shape_analysis:start
 - b. Download "fourier2.5_.jar"
 - c. Copy and paste the file into the Plugins folder of your ImageJ installation
- 4) Open ImageJ
- 5) You should see this window:



Extracting the markings on each animal

Now that the programmes are up and running, we can open the digital images of our leopards and try to extract the black markings that we're interested to measure. We first threshold our images (so that the markings are black and everything else is white), then we can ask the program to detect the markings automatically.

- 6) Download the images from: <http://ianzwchan.com/my-research/lsm4267>
- 7) Open A1.tif (drag and drop onto ImageJ)
- 8) Click: Image>Adjust>Color Threshold...
- 9) Choose these settings (Threshold color = B&W; Dark background unchecked):

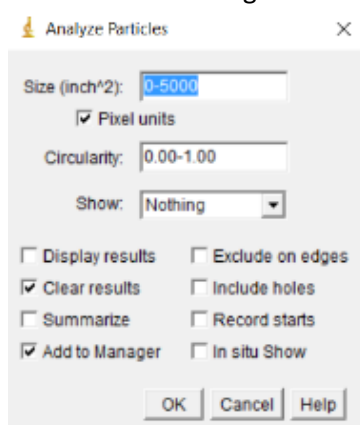


- 10) Click: Analyze>Analyze Particles...

Note that if you're trying to do this on one of the already-thresholded files (e.g.

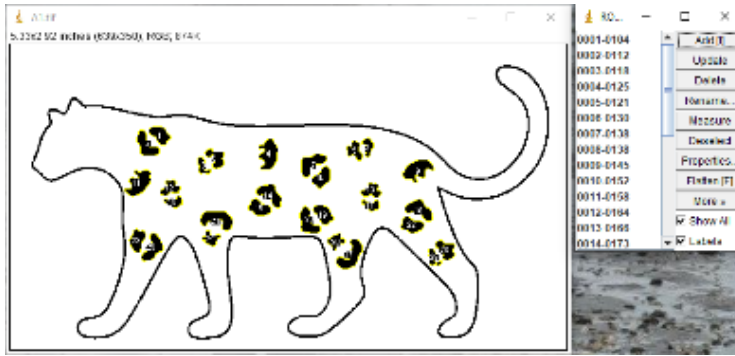
A2_thresholded.jpg), you first need to click: Image>Adjust>Threshold, and then "Apply"

- 11) Choose these settings:



- 12) Click: OK

13) You should now see something like this:



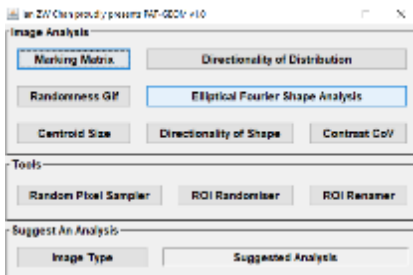
All the markings on the leopard should now be demarcated by yellow (or cyan, depending on your computer) lines, and there should be an ROI manager window with 29 numbers (Note: ROI = Region of Interest). This means we've successfully extracted the 29 markings on this leopard, and they can now be measured!

Measuring shape

Let's start with Shape. We measure shape using some fancy mathematical procedure called Elliptical Fourier Analysis. This produces some numbers (Descriptors) which quantitatively describe the marking's shape. Descriptor 1 describes its shape on a large scale and each successive Descriptor describes its shape at successively smaller scales.

14) Click: Plugins>PAT-GEOM v1.0.0>PAT-GEOM

15) You should see:



16) Click: "Elliptical Fourier Shape Analysis"

17) Click: "OK"

18) Wait awhile (windows will open and close rapidly)...

19) You should now see a "Results" window with a lot of numbers (29 x 20 to be precise):

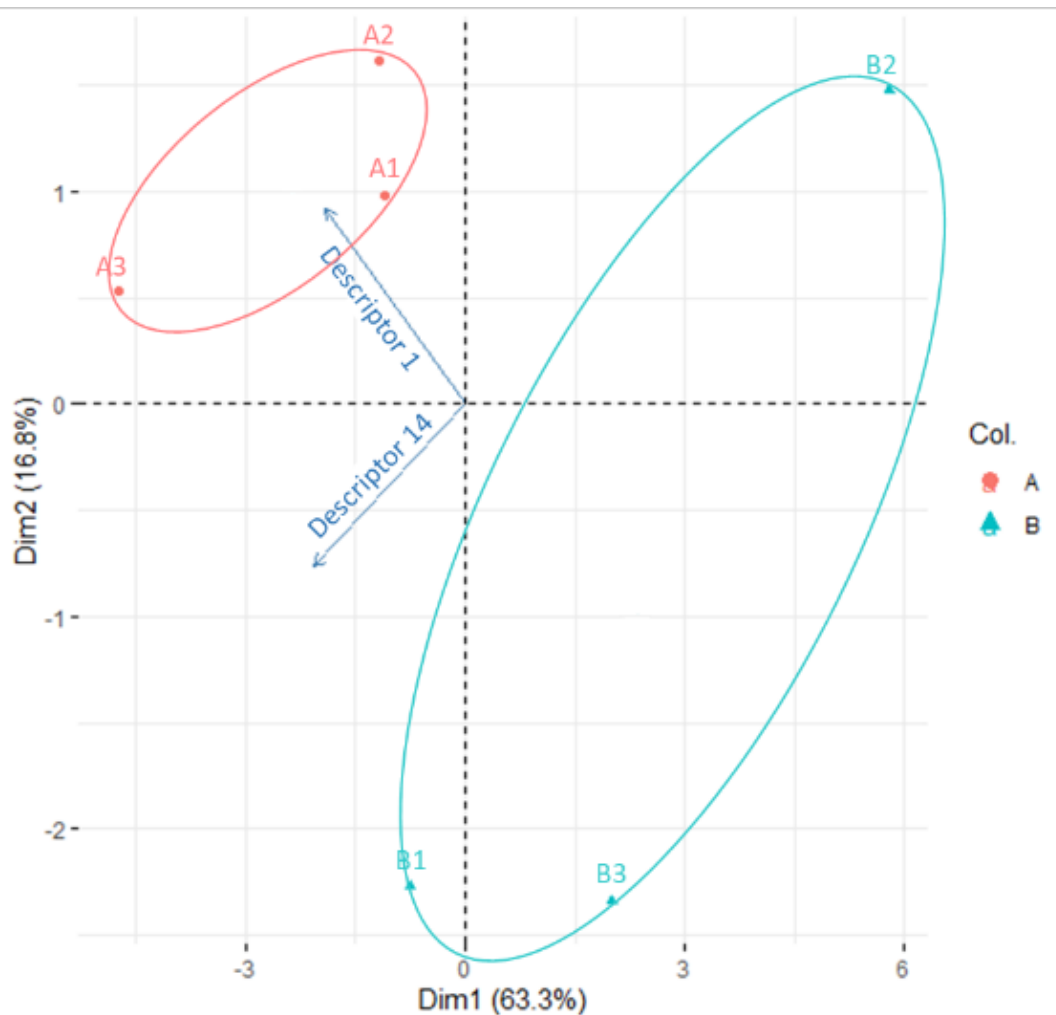
The screenshot shows a 'Results' window with a grid of 29 rows and 20 columns of numerical data. The columns are labeled with numbers 1 through 20. The rows are labeled with IDs such as '0001-0104', '0002-0112', etc., up to '0014-0173'. Each cell contains a numerical value representing a descriptor for that specific marking.

These are the results for Shape: 20 descriptors (rows) for each marking (columns). These can be copied (click on any cell and then click "Ctrl-A") and pasted into an excel file and organised. You could use these, for example, to get an idea of the "average marking shape" of each animal by

calculating the mean for each descriptor (Descriptor 1 to 20). These are the results for ONE leopard. We would have to do this for all the others as well (you're welcome to try if you want), but since this a class exercise let's just use the results I've provided in the "Data.xls" file.

One way to compare this data is to do a Principal Component Analysis (PCA). A PCA simply takes the 20 dimensions of the 20 means we calculated, reduces it, and plots it on a 2-dimensional axis for us to look at. On a PCA plot, points that are closer to each other are more similar and points that are further away from each other are more different.

This is what the PCA plot of mean marking shape looks like for our two leopard populations:



Above: Principal Component Analysis of the shape of the markings on the leopards from Population A (in pink) and Population B (in blue).

Do you notice any patterns or differences between the two populations?

Measuring size

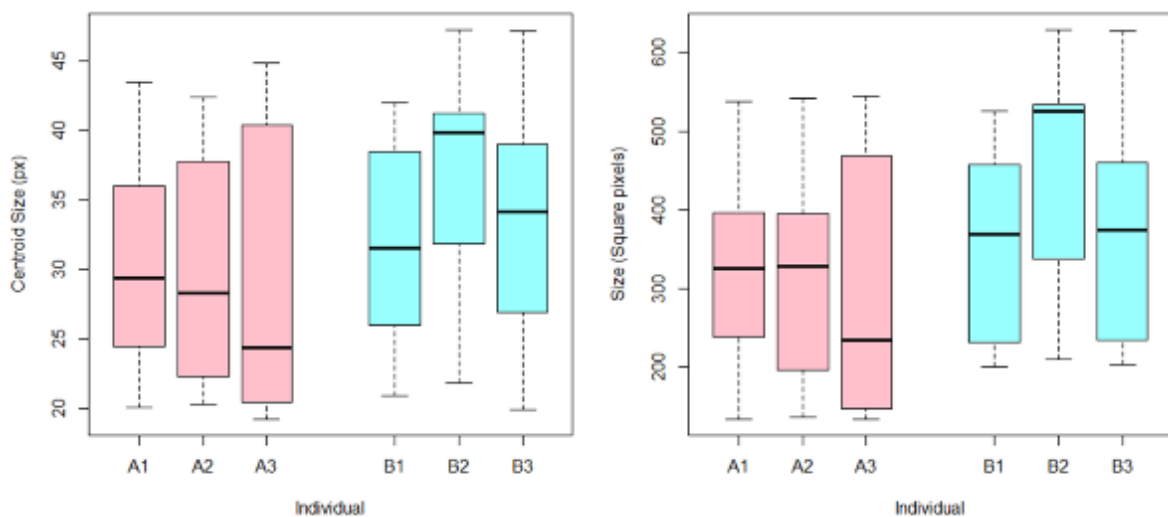
Next, Size. Here, we measure size in both another fancy mathematical formula that gives Centroid Size and the more intuitive Square Pixels.

- 20) Click back onto the PAT-GEOM window
- 21) Click: "Centroid Size"
- 22) The Results window should now look something like this:

Label	Average Centroid Size (px)	Size in Sq Pixels
1	30.315	332
2	34.338	352
3	32.350	348
4	39.317	534
5	36.367	481
6	30.797	209
7	36.221	470
8	30.726	183
9	40.819	527
10	39.885	727
11	33.832	339
12	35.289	347
13	35.300	233
14	36.181	234
15	30.738	184
16	29.229	225
17	31.468	147
18	37.007	326
19	29.878	228
20	36.169	186

This is the size of each marking on this leopard. Similarly, this could be used to calculate means to compare different animals and populations. Again, this would have to be done for each leopard but I've already done it and you can just refer to the other results in the "Data.xls" file.

Let's compare the sizes of the markings using boxplots of both size measurements:



Above: Boxplots of the size of the markings for the leopards from Population A (in pink) and Population B (in blue) quantified using Centroid Size (left) and Square Pixels (right).

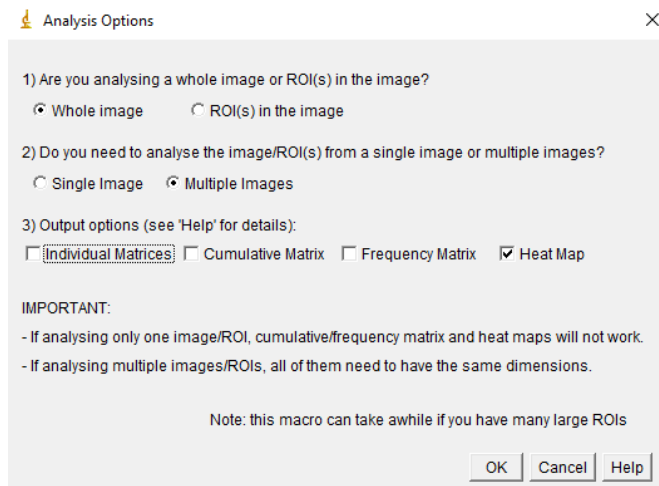
Do you notice any patterns or differences between the two populations?

Measuring distribution

Finally, let's measure distribution. The easiest way to inspect for any differences between the populations is to produce a Heat Map of where the markings are located for each population. With PAT-GEOM, this is easy.

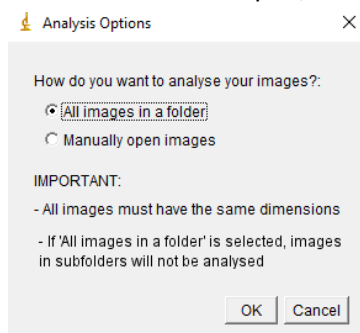
- 23) Click back onto the PAT-GEOM window
- 24) Click: "Marking Matrix"

25) Select these options:



26) Click: "OK"

27) A new window will open, select "All images in a folder"



28) Click: "OK"

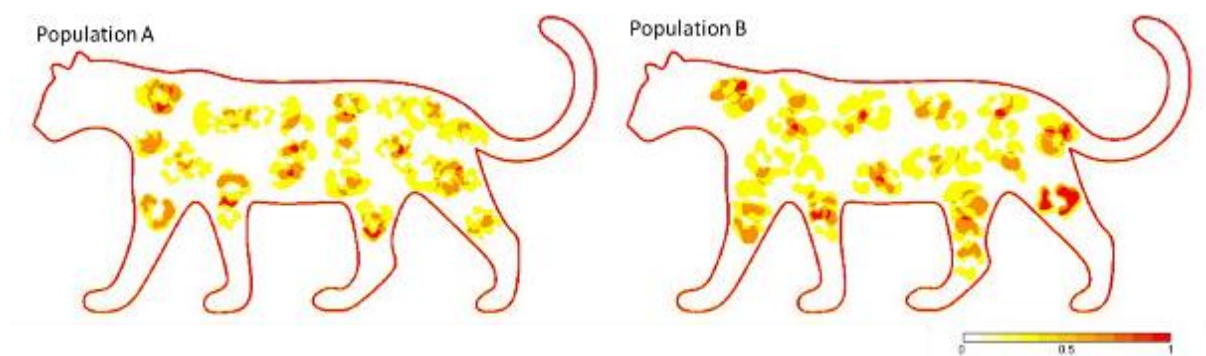
29) Select the "A_thresholded_inverted" folder and click "OK"

30) A Heat Map ("all_images_HeatMap.jpg") will now be saved in the "Pixel Matrix" folder in the "A_thresholded_inverted" folder

31) Repeat steps 23-30 for the "B_thresholded_inverted" folder

32) A Heat Map ("all_images_HeatMap.jpg") will now be saved in the "Pixel Matrix" folder in the "B_thresholded_inverted" folder

We can now compare these two Heat Maps (from Step 30 and Step 32):



Do you notice any differences between the two populations? Is there any characteristic that you could use to distinguish one population from the other?