

Exercise on PCA

November 12, 2021

The data set can be download from <https://archive.ics.uci.edu/ml/datasets/Anuran+Calls+%28MFCCs%29>.

1. Preprocess the data by centering the variables and divide them by their standard deviation.
2. Compute the Principal Components Analysis on these variables (You can use libraries for solving the eigenvalue problem.). Obtain and plot the eigenvalue spectrum using logscale for the y-axis.
3. Project the data in the two first PCs and color it by the family column.

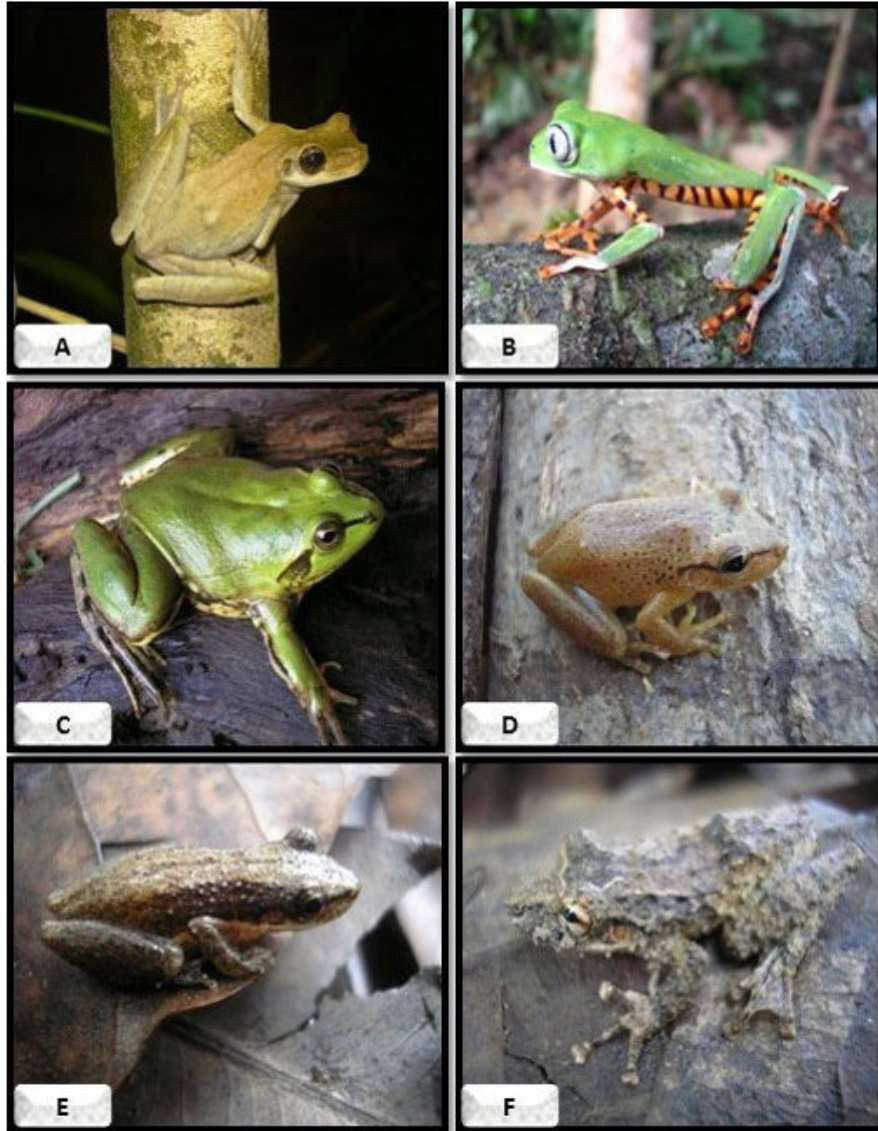


Figure 1: Some of the “data samples”