

---

## Advanced Biostatistics - ABSTAT17

IGC, 10th - 13th April, 2017

### Challenging group practical - PCA

Analyse the famous dataset on Painted Turtles (“tortues”), which can be found in the R package `ade4` (Jolicœur and Mosimann, 1960). “Long” corresponds to the length of the animal, “larg” to its width, “haut” to its height, “sexe” to its sex. All variables are expressed in centimeters with exception of the variable “sex”

1. Change the name of the variables accordingly with the up description.
2. Create a factor containing each individual’s sex and generate a colour vector to distinguish male and female.
3. Extract continuous variables in to a `data.frame`.
4. Start to make a descriptive analysis of your data.
5. Analyze correlation between variables (use scatter plots).
6. Do a PCA analysis with `princomp()` function.

### References

P. Jolicœur and J.E. Mosimann. Size and shape variation in the painted turtle. a principal component analysis. *Growth*, 24:339-354, 1960.