Projects Model Based Clustring:

- **Mixture Discriminant Analysis**: read Section 4.6 of the book of Charles Bouveyron available here (https://math.univ-cotedazur.fr/~cbouveyr/MBCbook/) and implement the EM algorithm from scratch.
- **Generate MNAR data**: explain why the MNAR mechanism is challenging (read Section 15.1 of the book of Little and Rubin (Statistical Analysis with Missing Data - send an email to Aude Sportisse to get the reference)), introduce synthetic MNAR values in a complete dataset and check that basic imputation methods (e.g. missforest) is not adapted in this case.
- **Comparison of LDA and QDA with BIC**: read Section 3 of the book Elements of Statistical Learning available here (https://hastie.su.domains/Papers/ESLII.pdf) on Quadratic Discriminant Alaysis (QDA), implement it from scratch and compare it with LDA (already implemented in class) by using Bayesian Information Criterion (BIC).
- **Model-Based Clustering of Count Data**: read Section 6.4 of the book of Charles Bouveyron available here (https://math.univ-cotedazur.fr/~cbouveyr/MBCbook/) and implement the Poisson Mixture Models from scratch.
- **SEM algorithm** read the paper here (https://inria.hal.science/inria-00074164/document) (and in particular the parts mentioning the SEM algorithm, implement the SEM algorithm for Gaussian mixtures and compare it with the EM algorithm for Gaussian mixtures (the EM algorithm will be implemented in class).