








Teaching methods	<p>The class consists of lectures (30h) and exercises sessions (7.5h). The classes and the TP are intended to be face to face. Teaching language: English.</p>
Content	<p>The class is focused on the presentation of key concepts of statistical learning and high-dimensional models such as:</p> <ul style="list-style-type: none"> • Statistical learning • Challenges concerning high-dimensional models and differences from low-dimensional models • Classical variable selection techniques for linear regression models: R^2, adj.R^2, C_p • Information criteria selection: KL divergence, AIC/TIC/BIC derivation • Cross-validation based selection: Leave-one-out and K-fold • Under- and overfitting or the bias-variance trade-off • Ridge shrinkage: theoretical properties, bias/variance trade-off, GCV • Lasso shrinkage: regularization paths, LARS, coordinate descent algorithm, prediction error bounds, degrees of freedom for lasso, support recovery, stability selection, knock-offs; inference by debiasing, post-selection inference, Bayesian inference • Extensions of Lasso: elastic net, group lasso, adaptive lasso, fused lasso • Other techniques: sparse graphical models and networks, sparse PCA, sparse Discriminant Analysis
Inline resources	<p>Moodle website of the class : LSTAT2450 - Statistical learning. Estimation, selection and inference. https://moodle.uclouvain.be/course/view.php?id=4214</p>
Bibliography	<ul style="list-style-type: none"> • Hastie, T., Tibshirani, R. and Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer. • James, G., Witten, D., Hastie, T., and Tibshirani, R. (2014). An Introduction to Statistical Learning: With Applications in R. Springer • Hastie, T., Tibshirani, R. and Wainwright, M. J. (2015). Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman and Hall/CRC. • Wainwright, M. J. (2019). High-Dimensional Statistics: A Non-Asymptotic Viewpoint. Cambridge University Press. • Bühlmann, P., van de Geer, S. (2011). Statistics for High-Dimensional Data. Springer.
Other infos	<p>Software: R/Python French friendly class.</p>
Faculty or entity in charge	<p>LSBA</p>

Programmes containing this learning unit (UE)				
Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Data Science : Statistic	DATS2M	5		
Master [120] in Statistics: Biostatistics	BSTA2M	5		
Master [120] in Mathematics	MATH2M	5		
Master [120] in Statistics: General	STAT2M	5		
Master [120] in Mathematical Engineering	MAP2M	5		
Master [120] in Data Science Engineering	DATE2M	5		
Certificat d'université : Statistique et science des données (15/30 crédits)	STAT2FC	5		
Master [120] in Data Science: Information Technology	DATI2M	5		