

## Scoring

Course title – Intitulé du cours	Scoring
Level / Semester – Niveau / semestre	M2 / first semester
School – Composante	Ecole d'Economie de Toulouse
Teacher – Enseignant responsable	Pascal Lavergne
Other teacher(s) – Autre(s) enseignant(s)	David Roux
Other teacher(s) – Autre(s) enseignant(s)	Arthur Mello
Other teacher(s) – Autre(s) enseignant(s)	
Other teacher(s) – Autre(s) enseignant(s)	
Other teacher(s) – Autre(s) enseignant(s)	
Lecture Hours – Volume Horaire CM	36 (Part 1 9H; Part 2 15H; Part 3 12H)
TA Hours – Volume horaire TD	
TP Hours – Volume horaire TP	
Course Language – Langue du cours	English / Anglais
TA and/or TP Language – Langue des TD et/ou TP	English / Anglais

### Teaching staff contacts – Coordonnées de l'équipe pédagogique :

1. Pascal Lavergne Office number: T506 [pascal.lavergne@ut-capitole.fr](mailto:pascal.lavergne@ut-capitole.fr)  
Preferred means of interaction: by appointment
2. David Roux [david.roux.consulting@gmail.com](mailto:david.roux.consulting@gmail.com)  
Preferred means of interaction: email
3. Arthur Mello [amello@avisia.fr](mailto:amello@avisia.fr)  
Preferred means of interaction: email

### Course Objectives – Objectifs du cours :

**Part 1 Theory (P. Lavergne)** Scoring goals and vocabulary. Introduction to the generalized linear model, and extensive study of logistic regression model, namely estimation and interpretation. Tools such as the Lorenz and ROC curves for the evaluation of the quality of a scoring model (and the selection of a score threshold for the logistic case). Codes are given in R.

**Part 2 Practice (D. Roux)** Banking Area is one of the very useful areas of application of scoring. Presentation of the main methodologies used. Practical application with the realization of a project per group of students by using SAS software.

**Part 3 Marketing applications (A. Mello)** Marketing is another important area of application for scoring. Presentation of marketing use cases for scoring: churn prediction, customer propensity to use a service, etc. Practical application with a project covering the entire data workflow using Python: data cleaning, feature engineering, modelling, data visualization, and presentation.

### Prerequisites – Pré requis :

Basic knowledge about linear regression and maximum likelihood. Practice of R, SAS, and Python softwares.

### **Practical information about the sessions – Modalités pratiques de gestion du cours :**

Personal computers allowed. The course attendance is compulsory

### **Grading system – Modalités d'évaluation:**

Part 1. Final exam on the theoretical part with exercises and interpretation of R code.

Part 2. Project per group of students - Bonus for participation. No late project accepted.

Part 3. Project per group of students - Bonus for participation. No late project accepted.

The final grade is the average mark of the three parts.

### **Bibliography/references – Bibliographie/références :**

#### *About generalized linear models*

Chapters 1 and 6 of Extending the linear model with R, J.J. Faraway, Chapman & Hall/CRC, 2006.  
Chapter 7 of W.N. Venables and B.D. Ripley, Modern Applied Statistics with S, 2002, Springer.  
Chapter 1 and 2 of Generalized additive models, an introduction with R, S. Wood, Chapman & Hall/CRC, 2006.

Chapter 2 of L. Fahrmeir and G. Tutz, Multivariate statistical modelling based on generalized linear models, Springer series in statistics, 1994. 2)

#### *About logistic regression and alternatives*

J.M. Hilbe, Logistic regression models, CRC Press, Chapman and Hall, 2009.

D.W. Hosmer, S. Lemeshow, Applied logistic regression, second edition, Wiley, 2000.

chapter 10 of S. Tufféry, Data mining et statistique décisionnelle, 2005.

S. Tufféry, Etude de cas en statistique décisionnelle, 2009.

chapter 2 of JP. Nakache and J. Confais, Statistique explicative appliquée, Technip, 2003.

M. Bardos, Analyse discriminante, application au risque et scoring financier, Dunod, 2001.

G. Shmueli et al., Data mining for business, techniques, and applications in R, Wiley, 2018.

#### *About the practice of scoring*

R. Anderson, The credit scoring toolkit, Oxford U.P., 2007. –

Thomas, Edelman and Crook, Credit scoring and its applications, SIAM, 2002.

N. Siddiqi, Credit risk scorecards, Wiley, 2006

### **Session planning – Planification des séances :**

**Part 1** To be announced

**Part 2**

Day	Duration	from	to
Tuesday 04/10/2022	03h00	09h30	12h30
Tuesday 04/10/2022	03h00	14h00	17h00
Monday 10/10/2022	03h00	09h30	12h30
Tuesday 11/10/2022	03h00	09h30	12h30
Tuesday 11/10/2022	03h00	14h00	17h00

### Part 3

Day	Duration	From	To
Tuesday 08/11/2022	03h00	14h00	17h00
Tuesday 15/11/2022	03h00	14h00	17h00
Tuesday 22/11/2022	03h00	14h00	17h00
Tuesday 29/11/2022	03h00	14h00	17h00

### **Distance learning – Enseignement à distance :**

Distance learning can be provided when necessary by implementing, for example: / En cas de nécessité, un enseignement à distance sera assuré en mobilisant, par exemple :

- Interactive virtual classrooms / Classe en ligne interactive
- MCQ tests and other online exercises and assignments / QCM et exercices en ligne
- Remote (online) tutorials (classes) / TP/TD à distance
- Chatrooms / Forums