

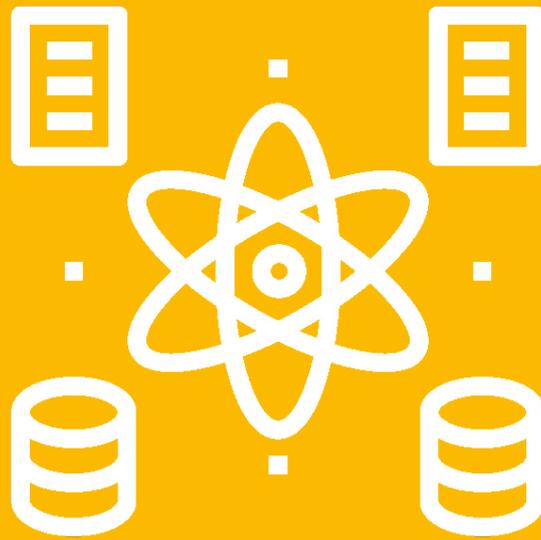
# Master 1

Sciences, Technologies, Santé

2025-2026

*Mathématiques et applications*

# Data Science



M1 DS

---

# SOMMAIRE

# CONTENUS

**03**

Contacts de la formation

**04**

Calendrier de la formation

**05**

Présentation de la formation

**08**

Volumes horaires et évaluations

**10**

Contenu des enseignements  
Semestre 1

**16**

Contenu des enseignements  
Semestre 2

**23**

Validation du M1 et poursuite d'études

Sommaire interactif  
pour revenir au sommaire  
cliquer sur 



---

# CONTACTS

Hélène TRICOÏRE-LEIGNEL : Directrice adjointe à la Pédagogie  
[helene.tricoire-leignel@univ-angers.fr](mailto:helene.tricoire-leignel@univ-angers.fr)

Gilles STUPFLER : M1 Responsable pédagogique et Président du Jury  
[gilles.stupfler@univ-angers.fr](mailto:gilles.stupfler@univ-angers.fr)

Sinan YALIN : Président du Jury  
[sinan.yalin@univ-angers.fr](mailto:sinan.yalin@univ-angers.fr)

Sandrine HERGUAÏS : Gestion de la scolarité et des examens  
Tél. : 02 41 73 54 85  
[sandrine.herguais@univ-angers.fr](mailto:sandrine.herguais@univ-angers.fr)

## Scolarité - Examens

Bâtiment A, Rez-de-chaussée

Horaires

8h30 > 12h30

13h30 > 16h30

Du lundi au vendredi

Fermé le mercredi après-midi



## Semestre 1

Rentrée et début des cours	Lundi 01 septembre 2025
Vacances d'automne 	Du samedi 25 octobre 2025 au dimanche 02 novembre 2025
Fin des cours du 1 <sup>er</sup> semestre	Vendredi 05 décembre 2025
Vacances de Fin d'année 	Samedi 20 décembre 2025 au dimanche 04 janvier 2026
Examens 1 <sup>er</sup> semestre - Session 1	Du mercredi 17 décembre 2025 au vendredi 19 décembre 2025 Du Lundi 05 janvier 2026 au mercredi 07 janvier 2026
Jury 1 <sup>er</sup> semestre - Session 1	Mercredi 28 janvier 2026

## Semestre 2

Début des cours	Jeudi 08 janvier 2026
Vacances d'hiver 	Du samedi 21 février 2026 au dimanche 01 mars 2026
Fin des cours du 2 <sup>ème</sup> semestre	Vendredi 17 avril 2026
Vacances de printemps 	Du samedi 18 avril 2026 au dimanche 26 avril 2026
Examens 2 <sup>ème</sup> semestre - Session 1	Du lundi 04 mai 2026 au mercredi 08 mai 2025
Pont de l'Ascension	Du Jeudi 14 mai 2026 au Dimanche 17 mai 2026
Soutenance de TER et projets	Du mercredi 20 mai 2026 au vendredi 22 mai 2026
Jury 2 <sup>ème</sup> semestre - Session 1	Mercredi 27 mai 2026
Examens 1 <sup>er</sup> et 2 <sup>ème</sup> semestre - Session 2	Du lundi 08 juin 2026 au vendredi 12 juin 2026
Jurys 1 <sup>er</sup> et 2 <sup>ème</sup> semestre - Session 2	Mardi 30 juin 2026

Planning susceptible de modifications



# PRÉSENTATION DE LA FORMATION

La **data Science** désigne une discipline à l'interface entre modélisation mathématique, statistique et informatique, née de la nécessité croissante de traiter et d'exploiter les données du big data. La data science est désormais l'outil essentiel d'aide à la décision dans des domaines d'activités extrêmement variés : banque, finance, assurance, e-commerce et grande distribution, communication et marketing, santé, agro-alimentaire, aéronautique et défense, internet des objets et télécoms, etc..

Le **parcours Data Science** (DS) du master mention Mathématiques et Applications est une formation de niveau un (Bac+5). Il vise à former des cadres data-scientists capables de mettre en œuvre les techniques mathématiques et informatiques inhérentes à ce métier, possédant en outre des compétences métiers propres aux secteurs d'activités visés par le choix d'options.

## Objectifs, compétences et capacités visés par la formation

Le master Data Science (DS) est une formation en deux ans qui débouche sur le diplôme national de Master mention Mathématiques et Applications, parcours Data Science. A l'issue de la formation, le ou la diplômé(e) possède :

**Des compétences en mathématiques.** Le diplômé maîtrise les méthodes de l'analyse statistique et algorithmique de l'information, du datamining et la classification, au machine learning et deep learning, ainsi que les méthodes associées d'optimisation. Le diplômé est en capacité de modéliser les processus aléatoires, de les simuler et de confronter le modèle aux données statistiques du problème.

**Des compétences en informatique.** Le diplômé du master dispose des compétences relatives aux outils classiques d'analyse des

données (R, Python,...), de data management (SGBDR et moteur de stockage SQL,...) et plus spécifiques aux données de grandes dimensions et/ou non structurées : MapReduce et environnement Hadoop/HDFS, entrepôts de données Hive, PiG et analyse de données, etc. L'aptitude à la visualisation des données de grande dimension, images et graphes complète ce volet de compétences.

**Des aptitudes professionnelles.** Le diplômé du master sait définir un projet, ses objectifs et son contexte, les modalités de réalisation et d'évaluation des actions, leurs priorités et leur planning. Il sait utiliser les technologies de l'information et de la communication, effectuer une recherche d'information à partir de sources diverses, organiser une veille technologique. Il a une bonne capacité de communication, une bonne maîtrise de l'anglais scientifique et technique (certification), et sait rédiger clairement une note ou un rapport selon des supports de communication adaptés.

**Des compétences métiers.** Suivant le choix d'option, le diplômé possède des aptitudes spécifiques au traitement et à l'analyse des données de la biologie (données omiques notamment), ou celles des entreprises dédiées aux technologies du numérique.

## Débouchés visés et rémunération des diplômés

Les diplômés du master Data Science (DS) occupent des emplois de cadre Data scientist (ou Data miner, Data analyst, etc..) couvrant un large spectre de secteurs d'activités, touchant aussi bien au domaine industriel qu'à celui du tertiaire. Sont plus particulièrement visés :

- Option données biologiques : secteurs de la santé, de la pharmacie, la bio-industrie,



l'agro-industrie, les laboratoires de type in-  
serm, Cirad, Institut Pasteur, CEA, etc .

- Option données numériques : tout sec-  
teur tel que sociétés de conseil en gestion  
des entreprises ou spécialisées en datami-  
ning, de la grande distribution, les grandes  
banques et compagnies d'assurances, etc.

Selon l'APEC, la rémunération des Data scien-  
tists se situe entre 32 et 40 keuros brut/ an  
en début de carrière. Les enquêtes insertion  
récentes montrent une excellente insertion  
professionnelle au niveau cadre, pour un  
salaire moyen brut mensuel de 2600 Euros  
(hors primes et 13ème mois).

Le master est essentiellement à visée pro-  
fessionnelle. Les très bons étudiants peuvent  
poursuivre par une thèse de doctorat, en  
particulier dans le cadre d'une convention  
de partenariat avec une entreprise ou dans  
contexte académique.



Exemples de stages ou d'emplois  
récents : Renault Di-RC, INRA, iAd-  
vize, CPAM Maine et Loire, LNE,  
Sodi-france-Netapsys, Manitou, Par-  
ker-Meggitt, SFR, Ineox, AVISIA, Metro Cash  
& Carry France, Tronico groupe Alcen,  
Imbrication, Institut de Cancérologie de  
l'Ouest, CHU-Angers, Axens IFP Technolo-  
gies, BRED Banque populaire, Groupe PSA,  
CT Ingénierie, Harmonie Mutuelle, Banque  
de France, VyV Care IT, Capgemini, So-  
pra-Steria, Sigma Informatique, Socomec  
SAS, Institut Pasteur, ANSM, GRTgaz, RTE,  
Thales, Factoryz, CNPTI,...

## Public visé, pré-requis, conditions d'admission, tarification

Le master Data Science est à capacité li-  
mitée et l'admission en première (M1-DS)  
et deuxième année (M2-DS) du master est  
sélective. En dehors des qualités acadé-  
miques, le recrutement accorde une part  
importante à la motivation des candidats.

**La validation du M1-DS entraîne l'admission  
de droit en M2-DS.**

**Formation initiale.** Le M1-DS s'adresse en  
priorité aux étudiants diplômés d'une licence  
de Mathématiques (parcours fondamental  
ou appliqué), d'une licence MiASHS validée  
avec aisance dans les matières mathé-  
matiques, d'une double licence maths/éco  
ou maths/info, ou de toute formation d'un  
niveau au moins équivalent. Sous réserve  
d'un niveau suffisant en mathématique, les  
dossiers d'étudiants ayant eu un parcours  
plus dirigé vers l'informatique et/ou l'ingé-  
nierie sont considérés avec intérêt.

**Formation continue.** Les candidatures re-  
levant de la formation continue sont les  
bienvenues et considérées avec la plus  
grande attention. Les tarifs sont variables et  
communiqués suite à l'étude du dossier. Un  
ensemble de financements est disponible  
(se renseigner auprès de la Direction de la  
Formation Continue).

Les pré-requis à l'entrée du M1-DS couvrent  
le programme classique d'une Licence de  
Mathématiques ou tout autre programme  
européen ou international équivalent. **Outre  
l'aptitude mathématique, une connais-  
sance et une pratique de base d'outils de  
programmation sont recommandées** (Py-  
thon, et/ou R par exemple). Le diplôme de  
master visant des compétences en anglais  
de niveau B2 du CECRL, un niveau minimal  
B1 dans cette langue est recommandé à  
l'entrée du M1-DS.

L'enseignement se faisant essentiellement  
en français, **un niveau minimal B2 certifié  
en français est exigé des étudiants inter-  
nationaux à l'entrée du M1-DS.**



## Alternance en master Data SCIENCE

Le M2-DS est ouvert à l'alternance, en apprentissage et contrat de professionnalisation. Contactez les responsables de la formation pour plus d'informations et connaître les entreprises qui nous ont accordé leur confiance. (Attention : certaines recrutent près de 8 mois à l'avance et exigent un stage préalable en fin de M1-DS). Les responsables de formation se réservent la possibilité de refuser l'alternance à tout étudiant n'ayant pas validé le 1<sup>er</sup> semestre.

## Des raisons de choisir le master Data Science (DS) de l'UFR

La qualité de la recherche française en mathématiques fondamentales et appliquées est mondialement reconnue : suivant les critères choisis, elle se place du premier au troisième rang mondial. Dans ce cadre, la recherche en mathématiques du laboratoire LAREMA de l'Université d'Angers (une soixantaine de membres dont une quarantaine de permanents et des doctorants participant aux enseignements), associé au CNRS et partenaire du Labex Lebesgue, se distingue par son excellence. Aussi, **les étudiants du Master DS ont l'assurance de bénéficier des connaissances et des innovations les plus récentes**, celles à même d'intéresser les entreprises dans leur stratégie compétitive.

L'Université d'Angers est reconnue par ses résultats en matière de taux de réussite et d'insertion professionnelle : sur ces deux points, les enquêtes ministérielles la situent aux toutes premières places des universités françaises. **Les étudiants du Master DS ont l'assurance d'un accompagnement performant de leurs études jusqu'à leur insertion professionnelle.**

Angers et son agglomération, ville étudiante par excellence, apparaît dans les enquêtes nationales comme étant l'une des villes les plus attractives dans ce domaine. En résumé, **les étudiants du Master DS ont l'assurance d'une qualité de vie et d'étude particulièrement propice à leur réussite et à leur**

insertion professionnelle.



# VOLUMES HORAIRES - ÉVALUATIONS

## SEMESTRE 1

30 ECTS

UE	Matières	Volumes horaires					ECTS	Coeff.	Contrôle des connaissances			
		CM	TD	TP	Auto. Et.	Tot.			1 <sup>ère</sup> session		2 <sup>ème</sup> session	Durée CT
									Assidus	D.A.		
<b>Bloc Mathématiques</b>												
1	Analyse Numérique Matricielle	16	12	12	-	40	5	5	CT 0,4	CT 0,66 TP 0,34	1 CT 0,66 TP 0,34	2h30
									CC 0,3			
									TP 0,3			
2	Optimisation Non-Linéaire	16	12	12	-	40	5	5	CT 0,66 CC 0,34	CT	CT	2h30
3	Statistique **	32	22	0	-	54	6	6	CT 0,66 CC 0,34	CT	CT	3h
4	Modélisation Stochastique I	24	12	12	-	48	6	6	CT 0,66 CC 0,34	CT	CT	3h
<b>Bloc Numérique et Informatique</b>												
5	POO et Traitement des Données en Python	12	0	36	-	48	6	6	CT 0,5 TP 0,5	CT 0,5 TP 0,5	1 CT 0,5 TP 0,5	2h
	Mise à Niveau Logiciels et Systèmes	0	0	10	-	10	0	0	-	-	-	-
<b>Bloc Culture d'Entreprise</b>												
6	Anglais I	0	0	16	-	16	2	2	CC	Oral 0,5 CT 0,5	Oral 2	1h20
	Droit et Données Numériques	4	0	0	-	4	0	0	P	P	P	-
<b>Total Semestre 1</b>		<b>104</b>	<b>58</b>	<b>98</b>		<b>260</b>	<b>30</b>	<b>30</b>				

\*\* Note plancher 8/20

1 Notation TP basée sur projet (rapport + soutenance), avec report de note si supérieure ou égale à 10.

2 Oral pouvant se transformer en Écrit en Session 2.



### Conditions de validation du semestre 1

Admise si moyenne  $\geq 10$  et UE3  $\geq 8$

CT = Contrôle Terminal  
CC = Contrôle Continu

P = Validation en Présentiel  
DA = Dispensé d'Assiduité

Auto. Et. = Autonomie Étudiant



## SEMESTRE 2

## 30 ECTS

UE	Matières	Volumes horaires					ECTS	Coef.	Contrôle des connaissances			
		CM	TD	TP	Auto. Et.	Tot.			1 <sup>ère</sup> session		2 <sup>ème</sup> session	Durée CT
									Assidus	D.A.		
<b>Bloc Mathématiques</b>												
1	Datamining et Classification**	20	12	16	-	48	5	5	CT 0,66 CC 0,34	CT	CT	3h
2	Modèles de Régression	18	10	16	-	44	4	4	CT 0,5 TP 0,5	CT	CT	3h
3	Modélisation Stochastique 2	16	12	12	-	40	4	4	CT 0,66 CC 0,34	CT	CT	3h
<b>Bloc Numérique et Informatique</b>												
4	Structuration des données **	20	0	36	-	56	6	6	Oral 0,5 CT 0,5	Oral 0,5 CT 0,5	Oral 0,5 <sup>1</sup> CT 0,5	30mn 2h
<b>Bloc Culture d'Entreprise</b>												
5	Communication Scientifique	0	0	3	-	3	0	0	P <sup>4</sup>	P <sup>4</sup>	-	-
	Anglais 2	0	0	16	-	16	2	2	CC	Oral 0,5 CT 0,5	Oral <sup>2</sup>	1h20
	Insertion Professionnelle	0	0	4	-	4	0	0	P <sup>4</sup>	P <sup>4</sup>	-	-
	Travail d'Étude et de Recherche	0	3	0	37	40	5	5	Oral <sup>3</sup>	Oral <sup>3</sup>	Oral <sup>3</sup>	45mn
	Stage Facultatif	0	0	0	-	0	0	0	-	-	-	-
<b>Bloc Métiers au choix de l'étudiant 1 parmi 2</b>												
6	Option 1 : Introduction à la Génomique	20	0	12	-	32	4	4	CC <sup>3</sup>	Oral <sup>2</sup>	Oral <sup>2</sup>	
	Option 2 : Datamining et Économie	14	12	0	-	26	4	4	TP (dossier)	TP (dossier)	Oral	15mn
<b>Total semestre 2</b>		<b>91</b>	<b>43</b>	<b>109</b>	<b>37</b>	<b>280</b>	<b>30</b>	<b>30</b>				
						<b>Total année</b>	<b>540</b>	<b>60</b>				

\*\* Note plancher 8/20

\*\*\* Rapport facultatif (selon décision du tuteur) pas d'évaluation

- 1 Notation TP basée sur projet (soutenance) avec report de la note d'oral si supérieure ou égale à 10.
- 2 Oral pouvant se transformer en Écrit en Session 2.
- 3 Notation basée sur projet (rapport + soutenance).
- 4 Pas d'évaluation mais présence obligatoire (peut entraîner une défaillance à l'UE en cas d'absence injustifiée)



### Conditions de validation du semestre 2

Admis-e si moyenne  $\geq 10$  et UE1  $\geq 8$  et UE4  $\geq 8$



### Conditions de validation de l'année

Admis-e si  $(S1+S2) / 2 \geq 10$  et notes plancher du S1 et S2 atteintes

CT = Contrôle Terminal  
CC = Contrôle Continu

P = Validation en Présentiel  
DA = Dispensé d'Assiduité

Auto. Et. = Autonomie Étudiant



# CONTENU DES ENSEIGNEMENTS

## SEMESTRE 1

Les noms des responsables et intervenants sont donnés sous réserve de modifications.

### BLOC MATHÉMATIQUES UE OBLIGATOIRES

UE1

#### ANALYSE NUMÉRIQUE MATRICIELLE

*Numerical Matrix Analysis*

Responsable **Eric Delabaere**  
eric.delabaere@univ-angers.fr

#### PRÉREQUIS

##### *Notions et contenus*

Algèbre linéaire et bilinéaire en dimension finie (licence mathématiques L3) ; analyse numérique (licence mathématiques L3) ; langage Python.

##### *Compétences et capacités*

Maîtriser les notions principales de l'algèbre linéaire en dimension finie : applications linéaires et matrices, image et noyau, rang, changement de base, valeurs et vecteurs propres, matrice adjointe ; produits scalaires, normes vectorielles et normes matricielles ; connaître les propriétés principales des matrices symétriques et hermitiennes ; connaître les rudiments de la programmation sous Python.

#### CONTENUS

Complexité d'un algorithme ; conditionnement d'une matrice ; rayon spectral ; systèmes linéaires, résolution directe : méthodes de Gauss, factorisation LU et PLU, méthode de Cholesky, méthode QR ; moindres carrés ; systèmes linéaires, résolution itérative : méthode de Jacobi, méthode de Gauss-Seidel ; décompositions en valeurs propres et en valeurs singulières (SVD), recherche des valeurs propres : méthode de Jacobi, méthode QR, méthode des puissances.

#### COMPÉTENCES

– Connaître les conditions d'application des méthodes suivantes de résolution directe de systèmes linéaires, savoir les expliquer et les mettre en œuvre pour des matrices de petites

tailles : méthodes de Gauss, factorisation LU et PLU, méthode de Cholesky, méthode QR.

– Connaître les conditions d'application des méthodes suivantes de résolution itérative de systèmes linéaires, savoir les expliquer et les mettre en œuvre pour des matrices de petites tailles, savoir analyser leur convergence : méthode de Jacobi, méthode de Gauss-Seidel.

– Connaître les conditions d'application des méthodes suivantes de décomposition en valeurs propres ou en valeurs singulières, savoir les expliquer et les mettre en œuvre pour des matrices de petites tailles, savoir analyser leur convergence : méthode des puissances, méthode de Jacobi, méthode QR.

– Savoir expliquer ou construire un script Python des algorithmes précédents, en proposer des améliorations dans certains cadres applicatifs. Connaître et savoir utiliser sous Python des bibliothèques de type numpy ou scipy.linalg.

– Dans des cas pratiques simples, savoir modéliser un problème menant à la résolution de systèmes linéaires, le traiter numériquement sous Python par application des résultats du cours, et être capable d'interpréter les résultats obtenus.



G. Allaire, S.M. Kaber, Algèbre linéaire numérique. Ellipses (2002).

G. Allaire, Analyse numérique et optimisation. Editions de l'Ecole Polytechnique, 2005.

G.H. Golub, C.F. Van Loan, Matrix Computation. The John Hopkins University Press, 1989.



## UE2

### OPTIMISATION NON-LINÉAIRE

Nonlinear Optimization

Responsable **Jean-Marc Labatte**

[jean-marc.labatte@univ-angers.fr](mailto:jean-marc.labatte@univ-angers.fr)

#### PRÉREQUIS

*Notions et contenus*

Calcul différentiel en dimension finie, analyse (licence mathématiques L3) ; algèbre linéaire en dimension finie (licence mathématiques L3) ; analyse numérique (licence L3) ; langage Python.

*Compétences et capacités*

Maîtriser le calcul de dérivées, la formule de Taylor au second ordre d'une fonction  $C^2(\mathbb{R}^n)$  (gradient, hessienne) ; connaître et savoir calculer les droites tangentes et les vecteurs normaux à une courbe plane, à une courbe de niveau ; connaître et savoir calculer les plans tangents et les vecteurs normaux à une surface plane, à une surface de niveau ; maîtriser le calcul matriciel et l'interprétation géométrique de l'espace des solutions d'un système linéaire ; connaître les propriétés principales des matrices symétriques réelles et des formes quadratiques ; maîtriser le produit scalaire dans  $\mathbb{R}^n$  ; connaître les rudiments de la programmation sous Python.

#### CONTENUS

Programmation non-linéaire ; fonctions convexes en une et plusieurs variables ; optimisation sans contraintes ; méthode de descente de gradient ; méthode utilisant la hessienne (basée sur la méthode de Newton-Raphson pour résoudre une équation non-linéaire) ; multiplicateurs de Lagrange ; optimisation avec contraintes larges ; méthode de Karush-Kuhn-Tucker ; méthode de pénalisation (du point intérieur).

#### COMPÉTENCES

- Connaître le comportement et la caractérisation des ensembles et des fonctions convexes.
- Pour un problème d'optimisation donné, savoir reconnaître son type (optimisation avec ou sans contraintes) et savoir choisir la méthode adaptée pour le résoudre parmi

les suivantes : multiplicateurs de Lagrange, méthode de Karush-Kuhn-Tucker, méthode de pénalisation (du point intérieur).

- Dans des cas simples, savoir résoudre complètement un problème d'optimisation par mise en œuvre des méthodes précédentes.

- Comprendre et savoir utiliser sous Python des algorithmes standards d'optimisation convexe.

- Savoir résoudre des problèmes pratiques d'optimisation en une dimension (optimisation sans ou avec utilisation de la hessienne) en utilisant le logiciel Python.



N. Lauritzen, Undergraduate convexity: From Fourier and Motzkin to Kuhn and Tucker. World Scientific (2013).

S.G. Nash, A. Sofer, Linear and nonlinear optimization. McGraw-Hill (1996).

## UE3

### STATISTIQUE (Note plancher 8/20)

Statistics

Responsable **Mikael Escobar-Bach**

[mikael.escobar-bach@univ-angers.fr](mailto:mikael.escobar-bach@univ-angers.fr)

#### PRÉREQUIS

*Notions et contenus*

Théorie de la mesure et intégration (licence mathématiques L3) ; calcul des probabilités (licence mathématiques L3) ; algèbre linéaire en dimension finie (licence mathématiques L3).

*Compétences et capacités*

- Savoir manipuler les opérateurs du calcul des probabilités (espérance, variance, covariance, etc.) pour les variables discrètes ou continues.

- Maîtriser la notion d'indépendance des variables aléatoires et avoir les bases concernant la manipulation des vecteurs aléatoires.

- Connaître les lois des grands nombres et le théorème central limite.

- Avoir les bases de l'algèbre linéaire de Licence.



## CONTENUS

Rappels de statistique descriptive ; modélisation statistique ; estimation ponctuelle ; propriétés des estimateurs ; information de Fisher ; estimation de variance minimale ; tests d'hypothèse par rapport des vraisemblances et par intervalles de confiance ; échantillons gaussiens ; introduction à la statistique bayésienne ; analyse des données (ACP et AFC).

## COMPÉTENCES

- Connaître et savoir utiliser le vocabulaire propre à la statistique.
- Savoir proposer un modèle statistique adapté à un échantillon et identifier son paramétrage.
- Savoir estimer les paramètres du modèle (par la méthode des moments ou par maximum de vraisemblance) et leur associer des intervalles de confiance le cas échéant.
- Savoir calculer et interpréter l'information de Fisher apportée par un modèle sur son paramètre.
- Savoir comparer différents estimateurs sur la base des critères usuels.
- Savoir construire les intervalles de confiance adaptés au cadre des échantillons gaussiens.
- Comprendre le principe des tests d'hypothèses, savoir utiliser les intervalles de confiance pour construire des tests exacts ou asymptotiques.
- Comprendre le principe des tests non paramétriques à travers l'exemple des tests du khi-deux ou celui du test de Kolmogorov-Smirnov.
- Comprendre le principe de l'estimation bayésienne à travers quelques exemples simples.
- Comprendre les enjeux de la réduction de dimension à travers l'analyse en composantes principales.
- Comprendre le principe de l'analyse des données à travers les méthodes usuelles (ACP et AFC) et l'interprétation à en donner.



G. Saporta, Probabilités, Analyse des données et Statistique. Technip, 3<sup>e</sup> édition révisée, 2011.

## UE4

### MODÉLISATION STOCHASTIQUE 1

Stochastic Modeling 1

Responsable **Fabien Panloup**

[fabien.panloup@univ-angers.fr](mailto:fabien.panloup@univ-angers.fr)

### PRÉREQUIS

*Notions et contenus*

Cours d'intégration (licence mathématiques L3) ; Cours de Probabilités (licence mathématiques L3) ; Langage Python.

*Compétences et capacités*

- Maîtriser les éléments principaux de la théorie de la mesure et de l'intégration.
- Maîtriser les bases de la théorie des probabilités : variables discrètes et continues, vecteurs aléatoires, indépendance, modes de convergence, Borel-Cantelli, loi forte des grands nombres.
- Connaître les rudiments de la programmation sous Python.

### CONTENUS

*Compléments de probabilités* : Convergence de suites de variables aléatoires (presque sûre, en probabilité, en loi), loi conditionnelle (cas discret ou continu), introduction à l'espérance conditionnelle.

*Chaînes de Markov* : Chaînes de Markov à espace d'états fini ou dénombrable, définition et propriétés élémentaires, classification des états, temps d'arrêt et propriété de Markov forte, récurrence et transience, lois invariantes, temps d'atteinte, convergence à l'équilibre, théorème ergodique. Application à des modèles classiques (marches aléatoires, ruine du joueur, Ehrenfest, TCP, Wright-Fisher, ...)

*Simulation* : Modélisation et simulation numérique d'une v.a. de loi classique donnée, d'une suite de v.a. indépendantes. Modélisation et simulation d'une chaîne de Markov, de sa loi invariante. Illustration des convergences p.s. et en probabilité et des théorèmes de convergence à l'aide des simulations : loi des grands nombres, théorème limite central, théorèmes ergodiques pour les chaînes de Markov. Méthodes de Monte Carlo, notions sur les vitesses de convergence et la réduction de variance. Mise en pratique avec Python.

## COMPÉTENCES

- Connaître les notions élémentaires sur la convergence de variables aléatoires, savoir calculer une loi conditionnelle ou une espérance conditionnelle dans des cadres standards.
- Connaître les notions élémentaires sur les chaînes de Markov à espace d'états fini ou dénombrable : récurrence (positive ou nulle), existence et unicité de mesure/probabilité invariante, convergence en loi, calcul de temps d'atteinte et d'absorption.
- Savoir simuler une variable aléatoire, une chaîne de Markov et sa loi invariante.
- Savoir illustrer les différents résultats de convergence de suites de variables aléatoires et de chaîne de Markov.
- Savoir simuler une variable aléatoire réelle via les méthodes d'inversion, de rejet, de changement de variable.
- Savoir estimer le coût de simulation et le temps de calcul.
- Etre en capacité d'utiliser la méthode de Monte-Carlo.
- Mettre en oeuvre les méthodes de réduction de variance par variable de contrôle ou variables antithétiques.
- Etre capable de mettre en oeuvre des calculs de volumes, de prix de produits financiers, d'estimateurs sur des situations pratiques.



Nicolas Bouleau, Probabilités de l'ingénieur. Variables aléatoires et simulation. Editions Hermann (2002).

James Norris, Markov Chains. Cambridge University Press (1998) révisée, 2011.

## BLOC OUTILS NUMÉRIQUES ET INFORMATIQUE - UE OBLIGATOIRES

### UE5

## OUTILS INFORMATIQUES

### PROGRAMMATION ORIENTÉE OBJET (POO) EN PYTHON

*Object-Oriented Programming (OOP) in Python*

Responsable **Jacquelin Charbonnel**  
[jacquelin.charbonnel@univ-angers.fr](mailto:jacquelin.charbonnel@univ-angers.fr)

### PRÉREQUIS

#### *Notions et contenus*

Langage Python ; module de mise à niveau ; cours de programmation et d'algorithmie de Licence.

#### *Compétences et capacités*

Connaître les rudiments de l'environnement et du langage Python ; maîtriser les principes algorithmiques pour analyser un problème et concevoir une solution.

### CONTENUS

- Développement logiciel : introduction au génie logiciel, environnement et outils de développement (git).
- Python : modules, espaces de nommage, environnements virtuels, containers (modules collections et itertools), manipulation de fichiers, etc.
- Bonnes pratiques du développement informatique
- Programmation objet : principe de la POO, mise en pratique en Python : classes, héritage, exceptions.
- Réalisation d'un projet avec soutenance orale et/ou rapport.

### COMPÉTENCES

- Acquérir une bonne maîtrise du langage Python et de ses principales bibliothèques.
- Connaître et savoir mettre en œuvre les principaux concepts de la POO.
- Savoir appliquer ces acquis dans le cadre d'un projet de taille significative et de fonctionnalités diverses.

## UE5

### MISE À NIVEAU LOGICIELS ET SYSTÈMES

Software Refresher Course

Responsables Gilles Stupfler

gilles.stupfler@univ-angers.fr

Jacquelin Charbonnel

jacquelin.charbonnel@univ-angers.fr

#### CONTENUS

- Mise à niveau R (6h) facultative.
- Système Unix/Linux et connaissance du réseau du laboratoire : organisation du système de fichiers, shell protocoles réseau (4h).

#### COMPÉTENCES

- Connaître les rudiments du langage R et de ses environnements de travail (ex : RStudio). Satisfaire les prérequis des modules de Modélisation Stochastique 1 et 2, de Datamining et Classification et de Modèles de Régression.
- Savoir utiliser le système Unix/Linux et utiliser les divers services réseau du laboratoire.

## BLOC CULTURE D'ENTREPRISE UE OBLIGATOIRES

### UE6

#### CULTURE D'ENTREPRISE

##### ANGLAIS 1

English 1

Responsable Philippe Torres

philippe.torres@univ-angers.fr

#### PRÉREQUIS

*Notions et contenus*

Les bases de la langue anglaise.

*Compétences et capacités*

Maîtriser le niveau B1 du CECRL (dit « d'utilisateur indépendant »).

#### CONTENUS

- Permettre aux étudiants de continuer à travailler les cinq compétences en langue - compréhension écrite et orale, expression écrite et orale, et interaction orale - à travers des supports authentiques (articles, documentaires, documents audio et vidéo d'internet, graphiques...) et des activités variées (exercices de compréhension, d'expression écrite, jeux de rôle, débats, présentations orales...).
- Étoffer les connaissances lexicales.
- Améliorer la prononciation et revoir certains points de langue le cas échéant.
- Valider à l'issue du 2ème voire 3ème semestre un niveau B2 du CECRL (Certification).

#### COMPÉTENCES

On vise, à l'issue du 2ème voire 3ème semestre du Master, le niveau de compétence B2 du Cadre Européen Commun de Référence pour les Langues (CECRL) qui est résumé comme suit : « Peut comprendre le contenu essentiel de sujets concrets ou abstraits dans un texte complexe, y compris une discussion technique dans sa spécialité. Peut communiquer avec un degré de spontanéité et d'aisance tel qu'une conversation avec un locuteur natif ne comporte de tension ni pour l'un ni pour l'autre. Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités.»



## SUPPORTS OU RÉFÉRENCES

Les ressources mises à disposition en auto-formation recouvrent entre autres : plateforme LANSAD UA; Ressources en ligne (UNT UOH par exemple) ; Plateforme Moodle.

UE6

## COMMUNICATION SCIENTIFIQUE (LATEX)

*Scientific Communication (LaTeX)*

Responsable **Gilles Stupfler**

[gilles.stupfler@univ-angers.fr](mailto:gilles.stupfler@univ-angers.fr)

## CONTENUS

Ce cours concerne l'apprentissage de LaTeX, outil de communication scientifique largement utilisé, en particulier dès le semestre 2 pour le rapport du TER. Une partie de cet enseignement est en autoformation en salle informatique. Il pourra avoir lieu au semestre 2, selon les disponibilités et les besoins.

## COMPÉTENCES

- Maîtrise des rudiments de LaTeX.
- Être capable de rédiger un rapport de nature scientifique (et plus précisément mathématique) sous LaTeX (ex : le rapport de TER).

## SUPPORTS OU RÉFÉRENCES

Cours et vidéos mises à disposition sur Moodle pour un travail préalable à produire.

UE6

## DROITS ET DONNÉES NUMÉRIQUES

*Legal aspect of data*

Responsable **Gilles Stupfler**

[gilles.stupfler@univ-angers.fr](mailto:gilles.stupfler@univ-angers.fr)

Intervenante **Ingrid Dumont**

[idumont@angerone.fr](mailto:idumont@angerone.fr)

## CONTENUS

Ce cours a pour objet, à travers des grands thèmes du Droit de l'univers numérique, d'aborder les problèmes que les étudiants pourront rencontrer durant leur vie professionnelle, à l'occasion de la mise en place ou de l'exploitation ou encore de l'administration d'un système informatique.



# SEMESTRE 2

Les noms des responsables et intervenants sont donnés sous réserve de modifications.

## BLOC MATHÉMATIQUES UE OBLIGATOIRES

UE1

### DATA MINING ET CLASSIFICATION

(note plancher de 8/20)

Datamining and Classification

Responsable **Jean-Marc Labatte**

[jean-marc.labatte@univ-angers.fr](mailto:jean-marc.labatte@univ-angers.fr)

### PRÉREQUIS

#### Notions et contenus

Théorie de la mesure et intégration (licence mathématiques L3) ; calcul des probabilités (licence mathématiques L3) ; algèbre linéaire en dimension finie (licence mathématiques L3) ; modules de statistiques et d'analyse numérique matricielle du premier semestre ; langages R et Python.

#### Compétences et capacités

Maîtriser les compétences enseignées dans le module de statistiques du semestre 1 : statistique descriptive, modélisation, analyse des données ; maîtriser les bases de l'algèbre linéaire et du calcul matriciel, notamment : changement de base, recherche des valeurs propres, des valeurs singulières, la projection orthogonale de  $\mathbb{R}^n$  ; avoir une connaissance minimale des langages R et Python.

### CONTENUS

Principaux concepts du datamining ; analyse des données (ex : ACM, AFM, MDS) ; classification supervisée (k-plus proches voisins, analyse discriminante, arbres de décision, SVM, etc.) ; courbes ROC et AUC ; classification non-supervisée (k-means, classifications hiérarchiques, DBSCAN, etc.) ; mise en pratique sous R et/ou Python.

### COMPÉTENCES

– Mettre en pratique l'analyse des données vue en statistiques au premier semestre (ACP, AFC) ainsi que les méthodes vues dans le module (ex : ACM, AFM, MDS) : interprétation et démonstration par l'exemple. Être capable

d'interpréter graphiquement les sorties fournies par les logiciels de statistique pour en déduire des conclusions pertinentes en analyse des données.

– Comprendre le principe de la classification supervisée par l'intermédiaire des méthodes usuelles (k-plus proches voisins, analyse discriminante, arbres de décision, SVM, etc.). Mettre en œuvre ces méthodes sur des exemples concrets.

– Connaître le principe des algorithmes simples de classification non supervisée (k-means, classifications hiérarchiques, DBSCAN, etc.). Comprendre par l'exemple la structure d'un modèle de mélange et son estimation par un algorithme EM. Mettre en œuvre ces méthodes sur des exemples concrets.

– Savoir utiliser R ou Python pour largement exploiter les méthodes étudiées dans le module, en particulier les méthodes de classification et de réduction de dimension.



T. Hastie, R. Tibshirani, J. Friedman, The element of Statistical Learning. Data mining, inference, and prediction. Second edition. Springer Series in Statistics. Springer, New York, 2009.

R. Garreta, G. Moncecchi, Learning scikit-learn : machine learning in python. Packt Publishing, 2013.

Cornillon P.A., Guyader A., Husson F., Jégou N., Josse J., Kloareg M., Matzner-Løber E., Rouvière L., Statistiques avec R. PU Rennes, 3<sup>e</sup> édition revue et augmentée, 2012.



## MODÈLES DE RÉGRESSION

Regression Models

Responsable Gilles Stupfler

[gilles.stupfler@univ-angers.fr](mailto:gilles.stupfler@univ-angers.fr)

### PRÉREQUIS

#### Notions et contenus

Théorie de la mesure et intégration (licence mathématiques L3) ; calcul des probabilités (licence mathématiques L3) ; algèbre linéaire en dimension finie (licence mathématiques L3) ; algèbre linéaire en dimension finie (licence mathématiques L3) ; module de statistique du premier semestre ; langages R et Python.

#### Compétences et capacités

Maîtriser les compétences enseignées dans le module de statistique du semestre 1 : statistique descriptive, modélisation, estimation ; maîtriser les bases de l'algèbre linéaire et du calcul matriciel, notamment : projection orthogonale de  $\mathbb{R}^n$  ; avoir une connaissance minimale des langages R et Python.

### CONTENUS

Régression linéaire (simple/multiple) : existence et unicité de l'estimateur OLS, théorème de Gauss-Markov, estimation de la variance des erreurs, inférence, sélection de modèle, prédiction, validation des hypothèses. Analyse de variance (1F/2F).

Régression logistique : équations de vraisemblance, normalité asymptotique, séparabilité des données. Inférence, sélection de modèle, prédiction.

Modèle linéaire généralisé : principe de la validation croisée et indicateurs de qualité (AIC, BIC, etc.).

### COMPÉTENCES

– Être capable de construire un modèle permettant d'expliquer une variable quantitative par une combinaison linéaire de régresseurs quantitatifs. Savoir estimer les paramètres de ce modèle par la méthode des moindres carrés et leur associer des tests de significativité sous les hypothèses adéquates.

– Savoir construire un modèle d'ANOVA à un facteur et deux facteurs, avec ou sans interactions. Être capable d'estimer les paramètres

de la modélisation. Savoir construire les tests de significativité associés à de tels modèles sous les hypothèses adéquates (comparaison multiple, homoscedasticité, correction de Bonferroni, influence des facteurs, des interactions, etc.).

– Savoir manipuler la régression logistique afin de faire le lien avec le module de classification.

– Connaître le cadre des modèles linéaires généralisés, et savoir globalement les techniques d'estimation qui leur sont associées.

– Illustrer toutes les régressions étudiées avec R et/ou Python sur des jeux de données réelles.

– Savoir valider des modèles à l'aide des principaux indicateurs (AIC, BIC, etc.). Maîtriser la stratégie de la validation croisée.



JM. Azaïs, JM. Bardet, *Modèle linéaire par l'exemple*. Dunod, 2<sup>e</sup> édition, (2012).

P.A. Cornillon, E. Matzner-Løber, *Régression avec R*. Springer Editions, 2010.

G. Saporta, *Probabilités, Analyse des données et Statistique*. Technip, 3<sup>ème</sup> édition révisée (2011).

## UE3

### MODÉLISATION STOCHASTIQUE 2

Stochastic Modeling 2

Responsable **Loïc Chaumont**

[loic.chaumont@univ-angers.fr](mailto:loic.chaumont@univ-angers.fr)

#### PRÉREQUIS

*Notions et contenus*

Cours d'intégration (licence mathématiques L3) ; cours de probabilités (licence mathématiques L3) ; modélisation stochastique du premier semestre ; bases du langage Python.

*Compétences et capacités*

Maîtriser l'ensemble des compétences requises pour le cours de Modélisation Stochastique I ; connaître les bases sur les chaînes de Markov en temps discret ; connaître les méthodes classiques de simulation de variables aléatoires ; avoir une connaissance et une pratique minimale du langage Python.

#### CONTENUS

*Chaînes de Markov en temps continu* : Définition et propriétés. Temps d'atteinte, temps d'absorption. Lois invariantes et convergence à l'équilibre. Exemples du processus de Poisson, du processus de Poisson composé, du processus de naissance et mort et des processus de files d'attente. Étude du cas général des processus de branchement. Applications à des modèles de croissance de population. Processus de Poisson composé avec drift et applications à des modèles de ruine.

*Simulations* : Simulations sous Python de chaînes de Markov en temps continu. Simulation de la loi invariante, du temps d'absorption. Modélisation de problèmes de ruine et de croissance de populations.

#### COMPÉTENCES

- Connaître les notions élémentaires sur les chaînes de Markov à espace d'état fini ou dénombrable.
- Savoir simuler une variable aléatoire, une chaîne de Markov et sa loi invariante.
- Savoir illustrer les différents résultats de convergence de suites de variables aléatoires et de chaîne de Markov.
- Savoir utiliser la méthode de Monte Carlo.
- Connaître la définition des chaînes de Mar-

kov et savoir créer un modèle en temps discret et en temps continu.

– Plus particulièrement, savoir modéliser un problème de ruine (risque, assurance) ou bien de croissance de population en utilisant les processus classiques : processus de Poisson composé (avec drift), processus de naissance et mort.

– Savoir simuler sous Python des chaînes de Markov en temps discret et en temps continu.



James Norris, Markov Chains. Cambridge University Press (1998).  
Robert P. Dorow, Introduction to stochastic processes with R. Wiley

(2017).

### BLOC OUTILS NUMÉRIQUES ET INFORMATIQUE - UE OBLIGATOIRES

## UE4

### STRUCTURATION DES DONNÉES

(note plancher de 8/20)

Responsable **Corentin Lothodé**

[corentin.lothode@univ-angers.fr](mailto:corentin.lothode@univ-angers.fr)

#### PRÉREQUIS

*Notions et contenus*

Module POO en Python ; partie systèmes du module de mise à niveau.

*Compétences et capacités*

Connaissance et pratique du langage Python et de son environnement ; maîtriser les principes algorithmiques pour analyser un problème et concevoir une solution.

#### CONTENUS

- Bases de données relationnelles : utiliser/installer un SGBD (ex : MariaDB ou postgresql) ; les ORM (mappings objet-relationnel) Python (ex : SQLAlchemy).
- Données textuelles : encodage, traitement sous Unix.
- Structuration et formats d'échange de données : XML, JSON, HD.
- Technologies web, développement d'application web (framework Django).
- Réalisation d'un projet avec soutenance orale et/ou rapport.



## COMPÉTENCES

- Connaître l'algèbre relationnelle, les bases de l'écriture des requêtes SQL.
- Connaître les rudiments de l'administration de bases de données (créer une base de données, donner des droits à certains utilisateurs, etc.)
- Maîtriser l'intégration de requêtes vers un SGBD dans un langage de programmation.
- Maîtriser l'interaction avec des sites Web dynamiques (consultation et publication de données Web).
- Maîtriser les commandes Unix pour traiter des données textuelles, connaître les rudiments
- Connaître les formats XML et JSON et savoir les exploiter.
- Savoir mettre en œuvre les connaissances et savoir-faire acquis sur une application en relation avec l'UE métiers ou/et le projet.

## BLOC CULTURE D'ENTREPRISE UE OBLIGATOIRES

UE5

### ANGLAIS 2

English 2

Responsable **Philippe Torres**

[philippe.torres@univ-angers.fr](mailto:philippe.torres@univ-angers.fr)

### PRÉREQUIS

*Notions et contenus*

Les bases de la langue anglaise.

*Compétences et capacités*

Maîtriser le niveau B1 du CECRL (dit « d'utilisateur indépendant »).

### CONTENUS

- Permettre aux étudiants de continuer à travailler les cinq compétences en langue - compréhension écrite et orale, expression écrite et orale, et interaction orale - à travers des supports authentiques (articles, documentaires, documents audio et vidéo d'internet, graphiques...) et des activités variées (exercices de compréhension, d'expression écrite, jeux de rôle, débats, présentations orales...).
- Étoffer les connaissances lexicales.
- Améliorer la prononciation et revoir certains points de langue le cas échéant.
- Valider à l'issue du 2ème voire 3ème se-

mestre un niveau B2 du CECRL (Certification).

## COMPÉTENCES

On vise, à l'issue du 2<sup>ème</sup> voire 3<sup>ème</sup> semestre du Master, le niveau de compétence B2 du Cadre Européen Commun de Référence pour les Langues (CECRL) qui est résumé comme suit : « Peut comprendre le contenu essentiel de sujets concrets ou abstraits dans un texte complexe, y compris une discussion technique dans sa spécialité. Peut communiquer avec un degré de spontanéité et d'aisance tel qu'une conversation avec un locuteur natif ne comporte de tension ni pour l'un ni pour l'autre. Peut s'exprimer de façon claire et détaillée sur une grande gamme de sujets, émettre un avis sur un sujet d'actualité et exposer les avantages et les inconvénients de différentes possibilités. »

## SUPPORTS OU RÉFÉRENCES

Les ressources mises à disposition en autoformation recouvrent entre autres : plateforme LANSAD UA; Ressources en ligne (UNT UOH par exemple) ; Plateforme Moodle.

UE5

### INSERTION PROFESSIONNELLE

*Professional Integration Program*

Responsable **Gilles Stupfler**

[gilles.stupfler@univ-angers.fr](mailto:gilles.stupfler@univ-angers.fr)

Intervenante **Chloé Maillat (SUIO-IP)**

[chloe.maillat@univ-angers.fr](mailto:chloe.maillat@univ-angers.fr)

### CONTENUS

Rédaction de CV et préparation aux entretiens d'embauche (stages, alternance, emplois).  
Découverte du milieu professionnel.



**TRAVAIL D'ÉTUDE ET DE RECHERCHE (TER)***Introduction to Research Methods*

Responsable Gilles stupfler

gilles.stupfler@univ-angers.fr

Intervenants Tuteurs

**PRÉREQUIS***Notions et contenus*

Contenus du MI-DS du S1 (et en cours au S2).

*Compétences et capacités*

Connaissance des bases du logiciel LaTeX ; ensemble des compétences scientifiques et d'outils informatiques du MI-DS.

**CONTENUS**

Le Travail Encadré de Recherche (TER) est un travail de 4-5 mois réalisé en laboratoire par l'étudiant ou par un binôme d'étudiants. Encadré par un enseignant tuteur, il donne lieu à la rédaction d'un rapport et à une soutenance orale pouvant être faite en anglais. Ce travail met en œuvre les connaissances théoriques acquises pendant l'année sur des problèmes concrets issus de questions intéressant le milieu professionnel. Les documents d'appui peuvent être des textes en anglais, ce qui participe de la pratique de l'anglais scientifique et technique. Dans certains cas, le sujet du TER peut être mis en relation avec le stage facultatif. On autorise la pédagogie inversée, le TER peut être l'occasion de présenter à la classe une méthode statistique évoluée et non étudiée en cours.

**COMPÉTENCES**

- Savoir mettre en œuvre les connaissances théoriques acquises en mathématiques appliquées de l'année du MI-DS sur des problèmes concrets.
- Savoir produire un travail d'expérimentation numérique, à l'aide de logiciels et langages *ad hoc*, Python, R, SQL, ... et rédiger un rapport écrit selon des supports de communication *ad hoc* et des contraintes définies.
- Savoir définir et respecter des modalités de réalisation, leurs priorités, leur planning.
- Savoir s'exprimer à l'oral dans un cadre défini et travailler en équipe.

**STAGE FACULTATIF***Optional Internship Program*

Responsable Gilles stupfler

gilles.stupfler@univ-angers.fr

**CONTENUS**

Il est possible de réaliser un stage facultatif. Il est obligatoirement conventionné et, en particulier, il est soumis à l'appréciation du responsable de la formation. Ce stage ne donne pas lieu à une notation et n'est affecté d'aucun ECTS. Le stage est placé en fin d'année académique (de fin mai à fin août), il est à l'initiative exclusive de l'étudiant qui souhaite valoriser son cursus académique. Ce type de stage est recommandé pour les étudiants souhaitant accomplir le M2 en alternance. Un stage à l'international d'étude et de perfectionnement en langue étrangère est aussi possible. De nombreuses possibilités existent, se renseigner (très tôt) sur l'ensemble des partenariats entre l'Université d'Angers et d'autres universités. Il sera demandé aux étudiants concernés de résumer en quelques lignes le déroulement du stage au responsable de la formation, mais ni rapport ni soutenance ne seront exigés.

**COMPÉTENCES**

Acquérir une première expérience professionnelle pour aborder le M2 dans les meilleures conditions, commencer à se faire connaître du milieu professionnel.

# BLOC MÉTIERS - UE OPTIONNELLES (UN CHOIX SUR DEUX)

## Option 1

UE6

### INTRODUCTION À LA GÉNOMIQUE

*Introduction to Genomics*

Responsable **Claudine Landès**

[claudine.landes@univ-angers.fr](mailto:claudine.landes@univ-angers.fr)

#### CONTENUS

Cet enseignement vise à donner les notions de base en biologie afin de comprendre les concepts de la biologie moderne nécessitant des développements mathématiques en modélisation et en statistique.

– Introduction aux notions de base en biologie cellulaire et moléculaire :

- > Composition moléculaire du vivant.
- > La cellule eucaryote et procaryote.
- > Reproduction sexuée et hérédité.
- > Des gènes aux protéines.

– La génomique et les données omiques :

> Introduction aux méthodes NGS (Next Generation Sequencing).

> Génomique : séquençage de génomes complets (NGS) ; prédiction de gènes ; annotation structurale des génomes ; annotation fonctionnelle des génomes ; génomique des populations ; génomique comparative ; phylogénomique.

> Transcriptomique et épigénomique : expression des gènes et régulation de leur expression.

#### COMPÉTENCES

A l'issue de la formation, les étudiants :

- Sauront ce qu'est une cellule et connaîtront les différences entre procaryotes / eucaryotes.
- Auront des notions de chimie du vivant : constitutions de l'ADN, de l'ARN et des protéines.
- Comprendront le passage de l'information génétique d'un individu à ses descendants.
- Comprendront les variations génétiques au sein d'une population.
- Sauront utiliser les variations génétiques entre populations afin d'inférer les relations évolutives.
- Comprendront la régulation de l'expression génétique au sein d'une cellule ou d'un tissu qui fait qu'un même génotype conduit à des phénotypes différents au sein d'un individu ou d'une population.

## Option 2

UE6

### DATA MINING ET ÉCONOMIE

*Datamining and Economics*

Responsable **Christophe Daniel**

[christophe.daniel@univ-angers.fr](mailto:christophe.daniel@univ-angers.fr)

#### PRÉREQUIS

*Notions et contenus*

UE de statistique du premier semestre, UE parallèle de datamining et classification du second semestre ; autoformation en économétrie.

*Compétences et capacités*

Avoir des notions de base en statistique, analyse des données et économétrie. Rudiments de R.

#### CONTENUS

Analyse en composantes principales, analyse factorielle des correspondances, analyse des correspondances multiples, analyse factorielle multiple, analyse sur données mixtes, classification hiérarchique, régressions sur composantes principales, Moindres Carrés Partiels.

Ce cours d'analyse de données avancée et de statistiques multivariées présente les méthodes dont l'objectif principal est de simplifier la complexité des bases de données statistiques en mettant en évidence les corrélations entre les variables, les ressemblances entre les individus, et en perdant le moins possible d'information expliquée. Un lien est ensuite fait avec les cours d'économétrie par la présentation (sur R et sur SAS) des modèles de régressions sur variables latentes (PCR et PLS) qui permettent notamment de corriger les biais de multi-colinéarité entre les variables.

#### COMPÉTENCES

- Méthodes d'analyses factorielles classiques (ACP, AFC et ACM).
- Méthode de classification hiérarchique ascendante.
- Méthodes plus complètes : Analyse Factorielle Multiple, Analyse Factorielle sur Données



Mixtes, etc.

– Imputation des données manquantes (Miss Mda).

– Régressions sur variables latentes : Régressions sur composantes principales et Moindres Carrées Partiels..



Husson, Lê et Pagès : Analyse de données avec R, presses Universitaires de Rennes, 2009.

Cornillon, Husson, Jégou, Matzner-Løber : Statistiques avec R, Presses universitaires de Rennes, 2010.

# VALIDATION DU MASTER 1 ET POURSUITE D'ÉTUDES

## VALIDATION DES SEMESTRES ET DE L'ANNÉE M1-DS

Les modalités d'évaluation des compétences des UE et EC ont été détaillées. En complément de ces évaluations, un ensemble de devoirs (travail personnel sur temps libre) peut être demandé, selon un calendrier distribué aux étudiants en début de semestre. Ils sont préparatoires aux contrôles continus ou/et terminaux. L'enseignant peut éventuellement intégrer ces notes dans la note de contrôle continu (CC) sous réserve d'en avoir averti au préalable les étudiants.

La note (sur 20) attribuée au semestre est la moyenne pondérée des notes obtenues par UE, les coefficients de pondération étant identiques aux ECTS. L'obtention de la moyenne (10 et plus) sur un semestre entraîne la validation de ce semestre et des 30 ECTS associés. Une note plancher peut être définie pour certaines UE. Dans ce cas, l'obtention d'une note strictement inférieure à cette note plancher entraîne la défaillance (DEF) pour le semestre concerné, sauf avis contraire du jury.

La note (sur 20) sur l'année de M1-DS est la moyenne des notes obtenues pour les semestres 1 et 2. L'obtention de la moyenne (10 et plus) sur l'année M1-DS entraîne la validation de l'année et des 60 ECTS associés.

## JURY – REDOUBLEMENT EN M1-DS

Il est organisé trois jurys :

- Un jury d'évaluation chargé de valider les résultats d'examen obtenus dans les unités du 1er semestre.
- Un jury de 1<sup>ère</sup> session (après les examens du second semestre). Les règles de compensation entre semestres s'appliquent dès ce jury pour la validation de la première année de master.
- Un jury de 2<sup>ème</sup> session.
- Suivant le règlement de l'Université d'Angers, le redoublement du M1-DS est soumis à l'avis du jury. De manière générale, au plus un redoublement sera accordé (soit deux inscriptions), sauf décision d'extension accordée par le jury pour raisons spécifiques (maladie, étudiants

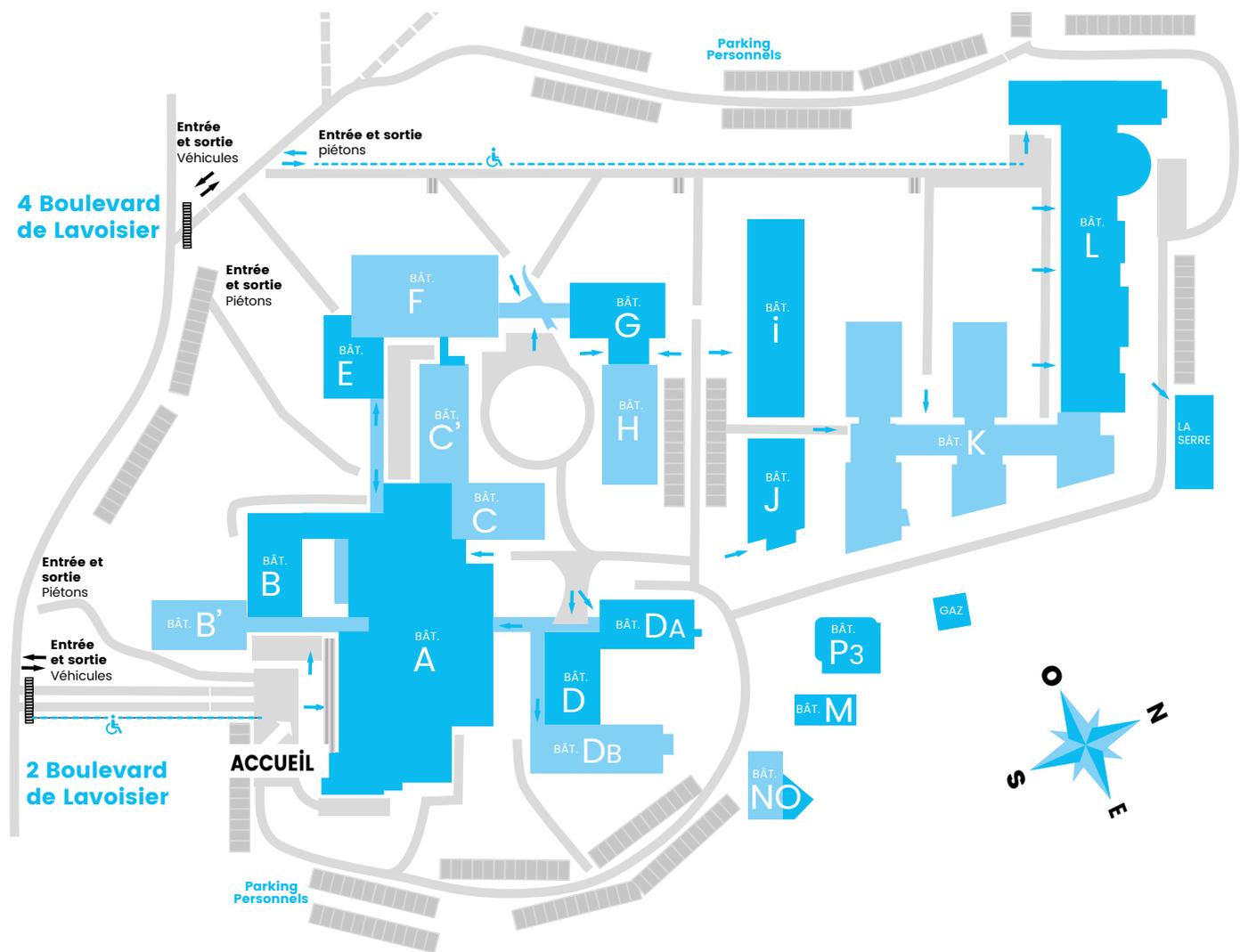
salariés, ...).

## POURSUIVE D'ÉTUDE EN M2-DS ANNÉE DE CÉSURE OBTENTION DU MASTER

La validation du M1-DS donne l'accès de droit en deuxième année M2-DS. En formation initiale et préalablement à leur accès en M2-DS, les étudiants ayant validé le M1-DS peuvent être autorisés à effectuer une année de césure, dans les conditions fixées par l'Université d'Angers et conformément à la circulaire MENS1515329C, n°2015-122 du 22-07-2015 du Ministère.

Le diplôme de Master est délivré à l'issue d'une année de Master 2 avec la validation de 120 ECTS sur les deux années.





- A** Scolarité | Accueil | Enseignement (Amphi A à E) | Administration
- B** Enseignement biologie
- B'** Enseignement biologie
- C** Enseignement chimie
- C'** Recherche
- D** Enseignement physique
- Da** Enseignement physique
- Db** Recherche
- E** Enseignement biologie
- F** Enseignement biologie | Recherche
- G** Enseignement géologie | informatique
- H** Enseignement informatique | Recherche
- i** Enseignement mathématiques | Recherche
- J** Enseignement chimie
- K** Recherche
- L** Enseignement transversaux | Enseignement (Amphi L001 à L006)

