

Pourquoi cette UE ?

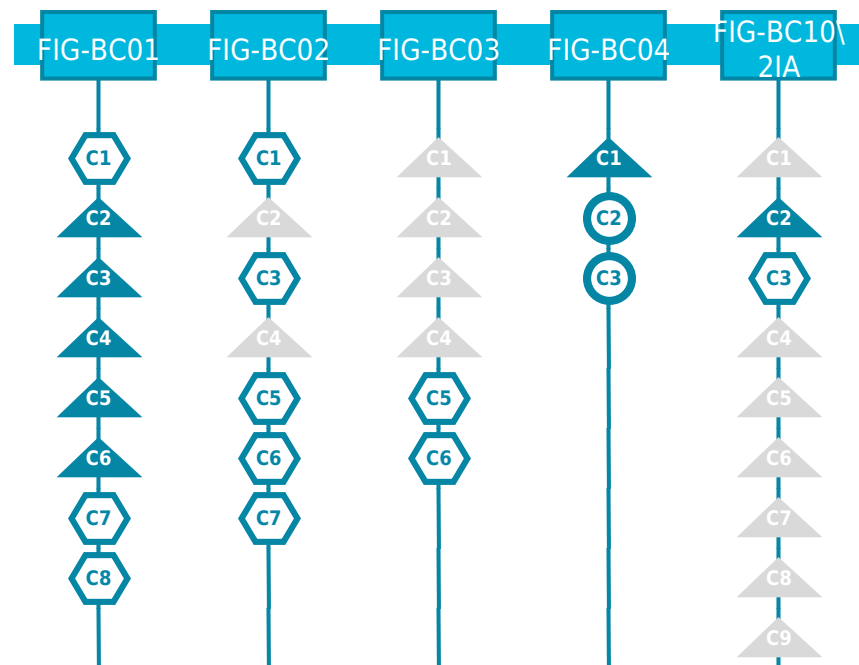
Ce module se focalise sur les données au sein d'un projet et plus particulièrement des étapes amont et aval d'un traitement informatique : leur acquisition (plan d'expérience, collecte) et leur visualisation (la partie gestion de base de données est, quant à elle, abordée dans le module 8.5). L'objectif est de sensibiliser les étudiants à l'importance de l'organisation de la collecte des données afin d'anticiper au maximum certains risques (échantillons trop petits, données imprécises, manquantes ou biaisées, informations redondantes, etc.).

Éléments constitutifs de l'UE

		coefficient
2IA_8_1-1 Collecte de données		1
2IA_8_1-2 Validation, Visualisation, Restitution		1
2IA_8_1-3 Projet de collecte, préparation et visualisation de données		2
Volume d'heures d'enseignement encadré	Volume d'heures de travail personnel	Nombre d'ECTS
59	16	4

Alignement curriculaire

Parmi les compétences visées par la formation, lesquelles sont développées dans cette UE ?



2IA_8_1 Collecte, préparation et visualisation de données	FIG
2IA_8_1-1 Collecte de données	S8

Contexte et enjeux de l'enseignement

Les enseignements du département 2IA, toutes options confondues, manipulent des données stockées de façon structurée ou non. La qualité des traitements informatiques sur ces données dépend de leur qualité. Cet enseignement pose les bonnes pratiques pour assurer cette qualité.

Prise en compte des dimensions socio-environnementales

Prérequis
 Régression linéaire, Python. Les étudiants doivent apporter leur ordinateur personnel à chaque séance.

Modalités d'enseignement et d'évaluation

	Nb d'heures
Cours	9
Cours intégré (cours + TD)	
TD	
TP	12
Projets	13
Travail en autonomie encadré	
Contrôles et soutenances	1
Travail personnel	10

Objectifs pédagogiques

(à la fin de cet enseignement, l'étudiant sera capable de ...)

Mettre en œuvre un plan d’expérience adapté.
Collecter et améliorer la qualité des données (utilisateurs, Web, capteurs).

Activités

(CM, TD, TP, projet, sortie terrain, etc.)

Les séances sont prévues pour 60 élèves et se déroulent sur ordinateur personnel. Organisation approximative :

Introduction (1h cours)
Plan d’expériences (1h cours, 1h TP)
Questionnaires (1h cours, 1h TP)
Web scraping (2h cours-TP)
Collecte via capteurs (6h cours-TP)
Qualité statistique des données (1h cours, 1h TP)
Qualité des formats (7h cours-TP)
Projet (10h séances, 3h soutenance)

Évaluations et retours faits aux élèves

(évaluations qui comptent pour la note ou qui permettent à l'étudiant de se situer, corrigés, feedback personnalisé...)

Contrôle des connaissances (QCM, coef. 1).
Les notions vues dans ce cours seront mises en application dans le projet global du module dont l'évaluation se fera au travers de rapport (coef 2) et de soutenances (coef 1) visant à éclaircir certains points du rapport.
Des évaluations non prévues à l’emploi du temps (contrôles surprise) peuvent avoir lieu.

3 semaines après la soutenance :
Retour des rapports et soutenances avec commentaires personnalisés.

2IA_8_1 Collecte, préparation et visualisation de données	FIG
2IA_8_1-1 Collecte de données	S8

Plan de cours

Plan d'expérience Questionnaires Web scraping et services REST Collecte à partir de capteurs Qualité statistique des données Qualité des formats de données (Pandas) Projet

Ressources et références

Diapositives et références.

2IA_8_1 Collecte, préparation et visualisation de données	FIG
2IA_8_1-2 Validation, Visualisation, Restitution	S8

Contexte et enjeux de l'enseignement

Les méthodes d’Intelligence Artificielle et de Science des Données sont toutes basées sur une compréhension de la donnée. Utiliser des méthodes de visualisations appropriées est une compétence indispensable pour mener ce travail à bien, en particulier savoir restituer les données via le bon type de graphique, transmettre les informations pertinentes de façon visuellement efficace, s’adapter à son public (néophyte ou expert), produire des rapports visuels et concis.

Prise en compte des dimensions socio-environnementales

Prérequis
 Connaissances basiques de R.

Modalités d'enseignement et d'évaluation

	Nb d'heures
Cours	3
Cours intégré (cours + TD)	
TD	
TP	3
Projets	
Travail en autonomie encadré	
Contrôles et soutenances	6
Travail personnel	6

Objectifs pédagogiques

(à la fin de cet enseignement, l'étudiant sera capable de ...)

Connaître les bonnes pratiques de visualisation et les types de graphiques adaptés.
Construire des graphiques avec R (ggplot2).
Intégrer ces visualisations dans un tableau de bord ou une application Web.

Activités

(CM, TD, TP, projet, sortie terrain, etc.)

Les enseignements sont prévus pour 60 élèves en promo complète. Les TP seront réalisés sur les ordinateurs personnels de ces derniers

Évaluations et retours faits aux élèves

(évaluations qui comptent pour la note ou qui permettent à l'étudiant de se situer, corrigés, feedback personnalisé...)

L'évaluation se fera par un projet : développement d'un tableau de bord (flexdashboard) ou d'une application R-Shiny mettant en jeu différents types de visualisation de données et à déployer sur le Web.

Retour sur l'évaluation fait à l'élève :
3 semaines après la dernière séance.

2IA_8_1 Collecte, préparation et visualisation de données	FIG
2IA_8_1-2 Validation, Visualisation, Restitution	S8

Plan de cours

Introduction à la visualisation de données (3h cours, 3h TP)
 ggplot2
 R Markdown
 Graphiques interactifs
 Cartographie
 Flexdashboard
 Applications Shiny
 Projet (6h TP).

Ressources et références

Supports de cours et références (livres, tutoriels).

Contexte et enjeux de l'enseignement

Dans le cadre de la transition numérique et des défis posés par la gestion des données, éventuellement massives, ce projet permet d'acquérir une compréhension pratique des outils et des méthodologies nécessaires pour garantir la qualité des données à chaque étape de leur traitement. En cohérence avec les principes de l'ingénierie responsable, il s'inscrit dans une démarche visant à soutenir des pratiques durables et fiables dans l'analyse des données.

Prise en compte des dimensions socio-environnementales

ODD9 - Industrie, innovation et infrastructure

Prérequis

Modalités d'enseignement et d'évaluation

	Nb d'heures
Cours	
Cours intégré (cours + TD)	
TD	
TP	
Projets	11
Travail en autonomie encadré	
Contrôles et soutenances	1
Travail personnel	

Objectifs pédagogiques

(à la fin de cet enseignement, l'étudiant sera capable de ...)

Le projet vise à familiariser les étudiants avec les processus d'acquisition et de visualisation des données, en mettant l'accent sur les étapes critiques situées en amont et en aval d'un traitement informatique, en particulier de Sciences des données. Les étudiants seront sensibilisés à l'importance d'une organisation rigoureuse de la collecte de données afin d'anticiper et de limiter les risques courants tels que les échantillons insuffisants, les données biaisées ou manquantes.

Compétences clés développées :

- Planifier et structurer un processus d'acquisition de données.
- Identifier et résoudre des problématiques liées à la qualité des données.
- Exploiter des outils de visualisation adaptés aux analyses complexes.
- Collaborer efficacement dans un cadre interdisciplinaire et partager des résultats clairs et pertinents.

Activités

(CM, TD, TP, projet, sortie terrain, etc.)

Projet

Évaluations et retours faits aux élèves

(évaluations qui comptent pour la note ou qui permettent à l'étudiant de se situer, corrigés, feedback personnalisé...)

Modalités d'évaluation :

- Soutenance orale avec présentation des visualisations et des conclusions critiques.
- Rapport de projet écrit évaluant la rigueur méthodologique et la pertinence des solutions proposées.

Grille d'évaluation intégrant des dimensions telles que la qualité de la réalisation, la capacité d'analyse et la clarté de la communication.

2IA_8_1 Collecte, préparation et visualisation de données	FIG
2IA_8_1-3 Projet de collecte, préparation et visualisation de données	S8

Plan de cours

Ressources et références