

## Master en Sciences de l'information

### Descriptif de module

**Domaine :** Economie et services

**Filière :** Information documentaire

**Intitulé de module** **Gestion des données** **2022-23**

**Code :** M6b

**Type de formation :**

Bachelor  Master  MAS  DAS  CAS  Autres : ...

**Niveau :**

Module obligatoire

Module spécialisé (à choix)

Module Travaux majeurs

**Domaine**

Sciences de l'information et des données

Recherche scientifique

Management

**Organisation temporelle :**

Module sur 1 semestre

Module sur 2 semestres

Semestre d'automne

Semestre de printemps

#### Prérequis

Avoir validé le module

Avoir suivi le module

Pas de prérequis

Autres : ...

#### Règle de validation du module

Module validé si moyenne des cours égale ou supérieure à 4

Module validé si moyenne des cours égale ou supérieure à 4 et unité de cours supérieure ou égale à 3

Autres : ...

**Crédits ECTS :** 6

#### Organisation – Descriptif par unité de cours

Code	Unité de cours	% du module
M6bC1	Data Curation	50
M6bC2	Data Visualisation	50

#### M6bC1 : Data Curation

Les données prennent une place fondamentale dans notre société, influençant les prises de décisions politiques, sanitaires et commerciales. La curation et la visualisation de données jouent un rôle principal dans le processus par lequel on extrait de la valeur des données brutes. Des bonnes représentations graphiques nous permettent, entre autres, de comprendre et extraire une information pertinente en identifiant les tendances dans les données et de communiquer le message clé à un public cible.

##### a) Objectifs visés :

- Comprendre et savoir expliquer les problématiques liées à la curation des données
- Comprendre et savoir expliquer les méthodes de curation des données
- Être capable de récupérer des données sous différents formats dans un format permettant la transformation
- Être capable de nettoyer, filtrer et enrichir les données récupérées

- Être capable de réaliser de la curation de données de manière collaborative et semi-automatisée
- Être capable de manipuler des données de différentes tailles (du méga au téra)
- Être capable de décrire et d'utiliser des formats de données standard pour la réutilisation et/ou l'archivage à long terme
- Être capable d'utiliser des outils ETL avancés ou en ligne de commande, et optionnellement par programmation
- Être capable de déposer ses données dans un dépôt de données ouverte

## b) Contenu et formes d'enseignement

Mots clés :

- Exploration, collecte, manipulation, filtrage, enrichissement, description, format, interopérabilité, traçabilité, réutilisabilité, automatisation, qualité des données.

Méthode d'enseignement :

- Cours théoriques, exercices et projet en partie en classe.

Langue principale d'enseignement :

- Français  Anglais  
 Autres : ..

## c) Modalité d'évaluation

Contrôle(s) continu(s) : 100% de l'unité de cours

- |   |  |  |
|---|--|--|
| <input type="checkbox"/> Evaluation écrite sur table                          | <input type="checkbox"/> Individuel            | <input type="checkbox"/> Groupe            |
| <input checked="" type="checkbox"/> Travail à rendre (30%)                    | <input checked="" type="checkbox"/> Individuel | <input type="checkbox"/> Groupe            |
| <input type="checkbox"/> TP   | <input type="checkbox"/> Individuel            | <input type="checkbox"/> Groupe            |
| <input checked="" type="checkbox"/> Oral groupé avec data visualisation (40%) | <input type="checkbox"/> Individuel            | <input checked="" type="checkbox"/> Groupe |
| <input type="checkbox"/> Autres : ...   | <input type="checkbox"/> Individuel            | <input type="checkbox"/> Groupe            |

## M6bC2 : Data Visualisation

Les outils de visualisation de données sont indispensables pour analyser d'énormes volumes d'informations, les comprendre et faciliter la prise de décision. En exploitant les capacités hautement visuelles de l'être humain, de bonnes représentations graphiques nous permettent, entre autres, de rapidement identifier les tendances dans les données et d'attirer le regard du public sur le message communiqué.

### a) Objectifs visés :

- Préparer des données brutes pour une visualisation correcte (par ex., filtrer, agréger, lier et/ou réunir des données) en utilisant R et *dplyr* ;
- Évaluer la pertinence d'une visualisation en tant que vecteur d'un message ;
- Choisir la visualisation adéquate en fonction du type de données et les objectifs visés ;
- Utiliser les bonnes stratégies visuelles pour transmettre un message-clé avec une visualisation ;
- Maîtriser R et *ggplot* pour créer des visualisations.

### b) Contenu et formes d'enseignement

Mots clés :

- Traitement de données, visualisation de données, RStudio.

Méthode d'enseignement :

- Chaque session comprend une partie théorique et une partie pratique. La partie pratique comprend des exercices et des travaux dirigés à réaliser dans l'environnement de travail RStudio. Des exercices de préparation aux sessions suivantes pourront être demandés aux étudiants. La révision du cours précédent est nécessaire avant chaque nouvelle session.

Langue principale d'enseignement :

- Français  Anglais  
 Autres : ..

## c) Modalité d'évaluation

Contrôle(s) continu(s) : 100% de l'unité de cours

- |  |                                     |  |
|--|-------------------------------------|--|
| <input type="checkbox"/> Evaluation écrite sur table                     | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe            |
| <input checked="" type="checkbox"/> Travail à rendre (30%)               | <input type="checkbox"/> Individuel | <input checked="" type="checkbox"/> Groupe |
| <input type="checkbox"/> TP  | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe            |
| <input checked="" type="checkbox"/> Oral groupé avec data curation (40%) | <input type="checkbox"/> Individuel | <input checked="" type="checkbox"/> Groupe |
| <input type="checkbox"/> Autres : ...                                    | <input type="checkbox"/> Individuel | <input type="checkbox"/> Groupe            |

## Modalités de remédiation et/ou répétition\* du module

- |   |   |
|---|---|
| <input checked="" type="checkbox"/> Remédiation possible entre 3.5 et 3.9                   | <input type="checkbox"/> Pas de remédiation   |
| <input type="checkbox"/> Répétition de l'ensemble des cours                                 | <input checked="" type="checkbox"/> Répétition des cours avec une note inférieure à 4 |
| <input checked="" type="checkbox"/> Répétition durant la session de rattrapage de septembre | <input type="checkbox"/> Répétition durant le semestre suivant                        |

\*Sauf exception, la répétition pour le Master IS se fait au niveau de l'évaluation et non du cours

## Remarques

(Facultatif – zone de rédaction libre)

Responsable de module : Arnaud Gaudinat

Descriptif validé par le responsable du module, le 17.02.2023

Descriptif validé par le responsable du Master IS, le 18.02.2023