



## DIPLÔME

- Diplôme d'Université  Diplôme Inter Universitaire

## PRÉSENTATION

INTITULÉ DU DIPLÔME : Big Data Data science et analyse des risques sous Python

DOMAINE :

- Droit, Économie, Gestion  Sciences, Technologies, Santé

Responsables : **Françoise Seyte, Stéphane Mussard**

[francoise.seyte@umontpellier.fr](mailto:francoise.seyte@umontpellier.fr) [stephane.mussard@umontpellier.fr](mailto:stephane.mussard@umontpellier.fr)

## ORGANISATION GÉNÉRALE

**Principales caractéristiques du projet** (*préciser le cas échéant, si le projet répond à une demande du monde socio-professionnel*) :

Projet créé pour les étudiants souhaitant approfondir leur connaissance en modélisation (analyse des risques) et les professionnels qui ont besoin d'une formation sur Python. La formation concerne : le big data, la data science, la manipulation de bases de données massives, le traitement des données (quantitatives, textes), la modélisation à haute dimension pour l'analyse des risques comme les réseaux de neurones et le machine learning.

**Objectif de la formation :**

Acquérir une formation concernant l'utilisation des bibliothèques de python avec notamment : pandas, sklearn, keras, tensorflow, pytorch, mongodb.

Compléter et enrichir la formation en VBA/SQL avec le NoSQL de python (Mongodb).

Obtenir une introduction à la programmation objet pour la programmation de réseaux de neurones discriminants nécessaires à l'analyse des risques.

Acquérir une formation (théorique et pratique) en détection de fraudes, et big data sous python, en passant par l'analyse juridique de l'utilisation des bases de données massives.

Acquérir les bases du web scraping afin d'extraire des informations (enrichissement de bases de données) et les analyser par les techniques de text mining et de machine learning.

**Compétences acquises au terme de la formation :**

Maîtriser python pour : les données massives (base de données : pandas, numpy, NoSQL), la modélisation (sklearn : séries temporelles, régressions, analyse des données, etc.), le text mining (extraire de la connaissance à partir de données textuelles), l'analyse des risques clients (tensorflow, keras, pytorch : réseaux de neurones discriminants), risque d'anomalies clients (détections de fraudes), python (et R pour l'actuariat).

**Débouchés professionnels visés :**

Data scientist, data analyst, risk manager (analyste des risques actuariels, bancaires, climatiques, industriels, marchés, sanitaires,...) et modélisations sous python.



Public concerné :

Formation initiale

Formation continue

Ouvert à l'apprentissage

Ouvert au contrat de professionnalisation

Droits d'inscription Formation initiale (cf grille tarifaire – Annexe 1)	Tarifs Formation Continue (à fixer en concertation avec le SFC)	Cadre réservé au service commun de la formation continue
550 €	2200 €	Visa SFC :

ORGANISATION PÉDAGOGIQUE

Nombre de session :  1  2

Enseignement			Préciser si obligatoire ou facultatif	Modalités pédagogiques (en heures)				Modalités de contrôle des connaissances					
Intitulé	Enseignement en langue étrangère	Enseignement dispensé à l'étranger		e-learning (non présentiel)	CM	TD	TP	Nature de l'épreuve			Durée de l'épreuve	Coef.	
	Oui / Non	Oui / Non	Ecrit					Oral	Contrôle continu / Terminal				
Introduction à Python	N	N	O		16								
Web scraping et services web	N	N	O		18					X		QCM 1H	1
Textmining	N	N	O		16								
VBA SQL / NoSQL	N	N	O		16								
Détection fraudes	N	N	O		8								
Gestion et traitement des données	N	N	O		16					X			1
Réseaux de neurones – Deep learning	N	N	O		20					X		QCM 1H	1
Big Data Assurance	N	N	O		16					X		QCM 1H	1
Analyse des données, Data viz & Econométrie	N	N	O		27					X			1
Machine Learning	N	N	O		25					X			1
Traitements d'images	N	N	O		12								
Programmation orientée objet	N	N	O		10								
Droit et Big Data	N	N	O		3								

VOLUME HORAIRE TOTAL

203 HEURES



<b>Stage</b> Un minimum de 200 heures d'enseignement dont 50 heures en présence des étudiants doit être requis pour réaliser un stage (hors médecine)- (Art. L612-8)	<input checked="" type="radio"/> Oui <input type="radio"/> Non	H
<b>Production d'un rapport</b>	<input checked="" type="radio"/> Oui <input type="radio"/> Non	
<b>Projet tuteuré</b>	<input checked="" type="radio"/> Oui <input type="radio"/> Non	105 H
<b>Présentation d'un mémoire</b>	<input type="radio"/> Oui <input checked="" type="radio"/> Non	

**Calendrier de la formation :**

**Rappel :** Conformément au calendrier universitaire voté annuellement en CFVU. Le bornage de l'année est fixé entre le 1<sup>er</sup> septembre et le 31 août pour l'année universitaire concernée.

Les examens et stages doivent avoir lieu avant la fin de l'année universitaire.

Début d'enseignement :	Octobre
Fin d'enseignement :	Février

## ÉQUIPE PÉDAGOGIQUE

**ENSEIGNANTS de l'établissement (liste obligatoire – Art L 613-2)**

Nom / Prénom	Corps/contrat	Composante	Enseignement
Gilles Michel	PRCE	UM	Web scraping / réseaux de neurones / VBA NoSQL / Machine learning
Stéphane MUSSARD	Professeur des Universités	Université de Nîmes	Text mining
Florian Olives	Banque	Crédit Agricole	Introduction à python
Alfred Mbaraidjim	Data scientist	Chef d'entreprise	détECTION fraudes / textmining / Machine learning
Yassine Laghzali	Actuaire	ALLIANZ	Big Data Assurance
Samuel Stocksieker	Actuaire	Université Aix Marseille	Econométrie