

**Certificat de compétence : Préparation et intégration des données pour l'IA  
Business & Decision University – Février 2020**

<b>REFERENTIEL DE COMPETENCES</b>	
<b>Activités et tâches</b>	<b>Compétences Associées</b>
Auditer la qualité des données	<ul style="list-style-type: none"> <li>Maîtriser la statistique descriptive monovariée</li> <li>Savoir utiliser la librairie Numpy en langage Python</li> <li>Savoir utiliser la librairie Pandas en langage Python</li> <li>Savoir réaliser un audit de qualité de données</li> <li>Connaitre les principales règles de RGPD</li> </ul>
Préparer, nettoyer et programmer le recodage des données	<ul style="list-style-type: none"> <li>Maîtriser la statistique bivariée</li> <li>Savoir tester et mesurer la paramétrie des données</li> <li>Savoir programmer en Python</li> <li>Savoir programmer le recodage des données pour les rendre paramétriques</li> <li>Savoir caractériser et traiter les valeurs manquantes</li> <li>Savoir identifier et traiter les valeurs aberrantes</li> <li>Maitriser le fonctionnement des bases de données SQL</li> <li>Pouvoir créer des jeux de données de test</li> </ul>
Programmer les indicateurs nécessaires au fonctionnement des algorithmes	<ul style="list-style-type: none"> <li>Savoir mettre en œuvre une stratégie inférentielle</li> <li>Maîtriser les bases de la statistique multivariée</li> <li>Être capable de mettre en œuvre du machine-learning non supervisé</li> <li>Être capable de mettre en œuvre du machine-learning supervisé</li> <li>Savoir utiliser les algorithmes du domaine public pour créer un modèle</li> </ul>

**CERTIFICAT de compétence « Préparation et intégration des données pour l'IA »  
Business & Decision University – Février 2020**

**REFERENTIEL DE CERTIFICATION**

**1. Modalités et déroulement de la session d'évaluation**

Les modalités d'évaluation de ce certificat s'articulent autour d'exercices pratiques, d'un questionnaire à choix multiples et d'un projet pédagogique :

- Les compétences les plus conceptuelles et théoriques font l'objet uniquement de questionnaires à choix multiples ;
- Les compétences les plus pratiques font l'objet :
  - o d'exercices qui permettent de valider que l'apprenant sait les mettre en application. Ainsi, ces compétences sont évaluées à la fois par des questionnaires à choix multiples et des travaux pratiques ;
  - o d'un projet pédagogique permettant de mettre en pratique et d'articuler l'ensemble des compétences du certificat

**- Exercices pratiques**

Les exercices pratiques sont réalisés pendant l'action de formation.

Le résultat attendu est la validation de la réalisation intégrale et complète de chaque exercice par le formateur. Finir l'exercice pratique permet de valider que la compétence évaluée est acquise

**- Questionnaire à choix multiples**

Les questions permettent à l'apprenant de valider qu'il a bien assimilé les compétences techniques.

Pour chaque question :

- 4 réponses possibles
- 1 seule bonne réponse

Obtenir un score supérieur à la moyenne permet de valider que la compétence évaluée est acquise

**- Un projet opérationnel pédagogique**

Au début de la formation, les candidats recevront une information sur le projet à réaliser, la constitution des groupes et le planning de réalisation du projet pédagogique. Le thème du projet sera communiqué pendant la formation lorsque les candidats auront abordé les compétences nécessaires à la réalisation du projet.

**CERTIFICAT de compétence « Préparation et intégration des données pour l'IA »  
Business & Decision University – Février 2020**

**2. Modalités et critères d'évaluation**

Compétences évaluées	QCM	Travaux pratiques	Projet pédagogique
<b>1. AUDITER LA QUALITE DES DONNEES</b>			OUI
1.1. Maîtriser la statistique descriptive monovariée	OUI		
1.2. Savoir utiliser la librairie Numpy en langage Python	OUI	OUI	
1.3. Savoir utiliser la librairie Pandas en langage Python	OUI	OUI	
1.4. Savoir réaliser un audit de qualité de données	OUI		
1.5. Connaître les principales règles de RGPD	OUI		
<b>2. PREPARER, NETTOYER ET PROGRAMMER LE RECODAGE DES DONNEES</b>			
2.1. Maîtriser la statistique bivariée	OUI		
2.2. Savoir tester et mesurer la paramétrie des données	OUI		
2.3. Savoir programmer en Python	OUI	OUI	
2.4. Savoir programmer le recodage des données pour les rendre paramétriques	OUI		
2.5. Savoir caractériser et traiter les valeurs manquantes	OUI	OUI	
2.6. Savoir identifier et traiter les valeurs aberrantes	OUI	OUI	
2.7. Maîtriser le fonctionnement des bases de données SQL	OUI	OUI	
2.8. Pouvoir créer des jeux de données de test	OUI		
<b>3. PROGRAMMER LES INDICATEURS NECESSAIRES AU FONCTIONNEMENT DES ALGORITHMES</b>			
3.1. Savoir mettre en œuvre une stratégie inférentielle	OUI	OUI	
3.2. Maîtriser les bases de la statistique multivariée	OUI		
3.3. Être capable de mettre en œuvre du machine-learning non supervisé	OUI		
3.4. Être capable de mettre en œuvre du machine-learning supervisé	OUI		
3.5. Savoir utiliser les algorithmes du domaine public pour créer un modèle	OUI		

**La certification est délivrée si les deux conditions suivantes sont satisfaites :**

- Une note globale égale à 70% de la note obtenue au QCM + 30% de la note obtenue au projet pédagogique supérieure à 10/20
- Accomplissement des exercices pratiques de chacun des modules