

Référentiel de compétences et de certification

<i>Intitulé de la certification</i>		
<i>Exploiter, analyser des données structurées et réaliser des études statistiques (Data Analyst)</i>		
<i>Description du métier, de l'activité ou de la situation professionnelle à partir duquel le dispositif de formation visant la certification est initié</i>		
<p>Le Data Analyst collecte les données, contrôle leur qualité et les représente de manière synthétique. De plus, il mène des analyses exploratoires et met en œuvre des modèles statistiques simples. Avec la certification « Exploiter, analyser des données structurées et réaliser des études statistiques (Data Analyst) », l'Ensaie-Ensaie Formation Continue (Cepe) met à la disposition des professionnels (chargés d'études, analystes, chefs de produit ...), ayant au moins trois ans d'expérience et un niveau en mathématiques au moins équivalent à un bac scientifique, un dispositif attestant des compétences nécessaires pour exercer le métier de data analyst.</p>		
<i>Référentiel de compétences</i>	<i>Référentiel de certification</i>	
	<i>Modalités d'évaluation</i>	<i>Critères d'évaluation</i>
C1 : Préparer ses données afin de les analyser efficacement en important et mettant en forme des jeux de données.	Mise en situation professionnelle par le biais d'un examen de 3 heures.	Cr1 : Le candidat a importé les données, il les a mises en forme, il a supprimé les doublons et identifié les valeurs erronées

<p>C2 : Traiter les données manquantes afin de compléter les informations non renseignées sans altérer significativement le jeu de données initiales en utilisant des techniques d’omission ou d’imputation.</p>	<p>Celui-ci se déroule sur ordinateur et comporte l’implémentation de 2 études de cas basées sur des données réelles.</p> <p>Après avoir importé et traité un jeu de données, le candidat en fera une analyse descriptive avancée et proposera un modèle de prévision simple.</p> <p>Le candidat rédigera un rapport écrit pour exposer, analyser et commenter les résultats obtenus et il joindra le code informatique ayant permis de parvenir à ces résultats.</p>	<p>Cr2 : Le candidat a détecté et imputé des données apparaissant comme manquantes dans les sources de données tout en conservant la cohérence de l’information initiale</p>
<p>C3 : Mener une analyse exploratoire d’un jeu de données afin de structurer et représenter l’information contenue dans les données en choisissant la méthode d’analyse factorielle adaptée (ACP, AFC ou ACM).</p>		<p>Cr3 : Le candidat a choisi une méthode d’analyse exploratoire appropriée aux données et à l’objectif. Il a su extraire de ces analyses les informations clés.</p>
<p>C4 : Etablir une typologie d’un ensemble de données (clustering) pour regrouper et classifier ces données en sous-ensembles appropriés partageant des caractéristiques communes, en sélectionnant un algorithme adapté au contexte (CAH et/ou K-means).</p>		<p>Cr4 : Le candidat a réduit la dimension du problème en mettant en place une classification dont il a su tirer des enseignements pertinents.</p>
<p>C5 : Réaliser des tests statistiques pour tirer des conclusions sur une population à partir d’un échantillon restreint en utilisant les techniques de test d’hypothèses.</p>		<p>Cr5 : Le candidat a correctement interprété les résultats du test statistique et en a restitué une synthèse explicite.</p>
<p>C6 : Mettre en oeuvre un modèle statistique pour obtenir des prévisions d’une variable d’intérêt à partir de facteurs explicatifs connus en utilisant les méthodes de régression linéaire ou logistique.</p>		<p>Cr6 : Le candidat a implémenté un modèle de prévision dont il a su questionner et évaluer la qualité de prédiction.</p>