

REFERENTIEL D'ACTIVITES <i>décrit les situations de travail et les activités exercées, les métiers ou emplois visés</i>	REFERENTIEL DE COMPETENCES <i>identifie les compétences et les connaissances, y compris transversales, qui découlent du référentiel d'activités</i>	REFERENTIEL D'EVALUATION <i>définit les critères et les modalités d'évaluation des acquis</i>	
		MODALITÉS D'ÉVALUATION	CRITÈRES D'ÉVALUATION
Bloc de compétences n°1 – créer une base de données et effectuer des requêtes afin de répondre à des besoins utilisateurs			

<p>Schématisme d'une base de données.</p> <p>Création de tables dans une base de données pour y charger des données.</p> <p>Création de requêtes dans une base de données.</p> <p>Mise à jour d'un catalogue de données.</p>	<p>Configurer une base de données afin d'y intégrer un jeu de données à exploiter.</p> <p>Effectuer des requêtes SQL pour répondre à une problématique métier.</p> <p>Mettre à jour un catalogue de données pour rendre accessible la base de données à ses utilisateurs.</p>	<p>Réalisation d'un projet (étude de cas) à partir d'un énoncé détaillé qui permet au candidat de se projeter dans une situation professionnelle reconstituée.</p> <p>Le projet donne lieu à une production de livrables par le candidat et à une soutenance devant un évaluateur.</p>	<p>Le candidat a créé et alimenté une base de données avec un outil de choix (MySQL, PostgreSQL, etc.)</p> <p>Le candidat a composé des requêtes simples (ex. SELECT, FROM, WHERE, JOIN, GROUP BY).</p> <p>Le candidat a composé des requêtes complexes (ex. de fenêtrage, les requêtes imbriquées, la clause HAVING).</p> <p>Le catalogue permet de récupérer automatiquement les informations décrivant les l'infrastructure des données (ex. fichier, base de données, table, colonne).</p> <p>Le catalogue offre des fonctionnalités de recherche et de filtrage.</p> <p>Le candidat a schématisé la configuration de la base de données et les tables ou relations (représentation visuelle ou sous la forme d'un ensemble de formules)</p>
--	---	--	--

--	--	--	--

Bloc de compétences n°2 – collecter et préparer des données de sources variées en utilisant un outil de gestion des données

<p>Collecte de données brutes de sources variées.</p> <p>Transfert des données collectées vers une zone de préparation de données.</p> <p>Préparation des données pour une exploitation ultérieure.</p>	<p>Collecter et intégrer des données venant de plusieurs sources en respectant les règles de protection des données en vigueur.</p> <p>Préparer des données pour l'analyse en respectant les normes internes à l'entreprise.</p>	<p>Réalisation d'un projet (étude de cas) à partir d'un énoncé détaillé qui permet au candidat de se projeter dans une situation professionnelle reconstituée.</p> <p>Le projet donne lieu à une production de livrables par le candidat et à une soutenance devant un évaluateur.</p>	<p>Le candidat a identifié plusieurs sources de données (base de données existantes, logs d'activité, les sites web, etc.) et en a extrait les données.</p> <p>Le candidat s'est approprié les règles internes de l'entreprise (concernant par exemple, le format des données, leur mode de stockage) et les applique aux données brutes.</p> <p>Les doublons ont été supprimés.</p>
---	--	--	--

Bloc de compétences n°3 – nettoyer et décrire un jeu de données à l'aide d'un langage de programmation spécialisé afin d'en préparer l'exploitation

Classification des différents types de données. Correction des erreurs et des incohérences présentes sur les données stockées. Réalisation d'une analyse uni-variée des données.	Prétraiter des données stockées en les classifiant et en corrigeant les erreurs et les incohérences présentes. Réaliser une analyse uni-variée de données prétraitées afin de les décrire et de détecter les incohérences.	Réalisation d'un projet (étude de cas) à partir d'un énoncé détaillé qui permet au candidat de se projeter dans une situation professionnelle reconstituée. Le projet donne lieu à une production de livrables par le candidat et à une soutenance devant un évaluateur.	Le candidat a repéré les différents types d'erreurs (les valeurs manquantes, les doublons et les outliers sont traités). Le code met correctement en œuvre les méthodes appropriées pour traiter les données. Le code contient au moins une opération de nettoyage (champs erronée, données manquantes, etc.)
--	---	---	---

Bloc de compétences n°4 – formaliser les résultats d'une analyse statistique en utilisant des outils de visualisation

<p>Analyse et problématisation d'un besoin client.</p> <p>Définition des indicateurs de suivi.</p> <p>Réalisation de graphiques.</p> <p>Création d'un tableau de bord.</p>	<p>Analyser un besoin client pour formuler des questions analytiques et définir des indicateurs de suivi.</p> <p>Créer les graphiques permettant de représenter les résultats obtenus.</p> <p>Créer un tableau de bord (dashboard) permettant de visualiser et de mesurer les résultats obtenus en fonction d'indicateurs donnés.</p>	<p>Réalisation d'un projet (étude de cas) à partir d'un énoncé détaillé qui permet au candidat de se projeter dans une situation professionnelle reconstituée.</p> <p>Le projet donne lieu à une production de livrables par le candidat et à une soutenance devant un évaluateur.</p>	<p>Le candidat a transformé une question "business" en question analytique.</p> <p>Le candidat a défini les indicateurs de performance pertinents (KPI).</p> <p>Le tableau de bord est lisible et répond aux besoins du client.</p> <p>Le tableau de bord est dynamique.</p> <p>Les formats de visualisation choisis sont adaptés à la problématique métier (ex. boxplot, barplot, pie chart, etc.)</p>
<p>Bloc de compétences n°5 – réaliser l'analyse exploratoire de données structurées afin de les synthétiser et de les interpréter</p>			

<p>Conduite d'analyses multivariées de données.</p> <p>Interprétation de séries temporelles.</p> <p>Réalisation d'un test statistique.</p>	<p>Réaliser une analyse multivariée pour comprendre la corrélation entre des variables et les synthétiser.</p> <p>Analyser des séries temporelles pour mesurer un phénomène au cours du temps.</p> <p>Réaliser un test statistique afin de vérifier une prédiction.</p>	<p>Réalisation d'un projet (étude de cas) à partir d'un énoncé détaillé qui permet au candidat de se projeter dans une situation professionnelle reconstituée.</p> <p>Le projet donne lieu à une production de livrables par le candidat et à une soutenance devant un évaluateur.</p>	<p>Les méthodes d'analyse statistique mobilisées sont pertinentes au regard des données à disposition et des objectifs de l'analyse.</p> <p>Le code permettant l'analyse est correctement effectué. Les résultats de l'analyse sont cohérents.</p> <p>Les résultats de l'analyse sont représentés graphiquement et formalisés de façon synthétique.</p>
<p>Bloc de compétences n°6 – réaliser une modélisation statistique à l'aide de méthodes d'apprentissage supervisé et non supervisé</p>			

<p>Réalisation d'une analyse prédictive à l'appui d'une méthode d'apprentissage supervisé (ex. régression linéaire, régression linéaire généralisée).</p> <p>Réalisation de classifications à l'appui d'une méthode d'apprentissage non supervisé.</p>	<p>Mettre en œuvre une méthode d'apprentissage supervisée pour réaliser une analyse prédictive.</p> <p>Mettre en œuvre une méthode d'apprentissage non supervisée pour opérer des classifications automatiques et partitionner des données.</p>	<p>Réalisation d'un projet (étude de cas) à partir d'un énoncé détaillé qui permet au candidat de se projeter dans une situation professionnelle reconstituée.</p> <p>Le projet donne lieu à une production de livrables par le candidat et à une soutenance devant un évaluateur.</p>	<p>Les méthodes mobilisées permettent de répondre à la problématique du client.</p> <p>Le code met en œuvre les algorithmes d'apprentissage supervisé et non supervisé correspondants.</p> <p>Les résultats de la modélisation sont cohérents.</p>
--	---	--	--