

Auto-évaluation

P11 - Data Analyst- Produisez une étude de marché avec R ou Python



Un dernier doute avant l'envoi de vos livrables ?

Pour vérifier la qualité de votre travail :

- cochez les cases ci-dessous : elles indiquent que vous avez bien pris en compte chaque indicateur de réussite ;
- renseignez, si besoin, la colonne "Notes" avec des commentaires sur vos livrables / vos étapes. Ils seront des points de discussion avec votre mentor pendant votre session de bilan / soutenance.

Quand toutes les cases de ce document seront cochées, vous pourrez déposer vos livrables sur la plateforme.

Bonne réussite !

Compétences	Livrables	Indicateurs de réussite de l'activité	Notes
Sélectionner les variables pertinentes d'un modèle d'apprentissage an de réaliser un apprentissage optimal	Le notebook ou chier R contenant la préparation et le nettoyage des données	<input type="checkbox"/> J'ai choisi au minimum 8 variables dans mon modèle. <input type="checkbox"/> Mon analyse comprend au minimum 100 pays. <input type="checkbox"/> J'ai vérifié : les données de toutes les variables sont complètes pour l'ensemble des pays. Je n'ai pas de vide. <input type="checkbox"/> Je suis satisfait de ma description des données : elle inclut une compréhension des outliers. <input type="checkbox"/> Je suis en mesure de justifier la pertinence de chaque variable. <input type="checkbox"/> J'ai créé au moins 3 nouvelles variables via du feature engineering.	
Réaliser des analyses multivariées pour comprendre la corrélation entre les variables	Le notebook ou fichier R de l'analyse	<input type="checkbox"/> Je suis satisfait de la réalisation de mon ACP. <input type="checkbox"/> J'ai mis en valeur les corrélations entre les variables.	

Réduire la dimension des jeux de données an d'optimiser les temps d'apprentissage des modèles	Le notebook ou fichier R de l'analyse	<input type="checkbox"/> J'ai réalisé un centrage des données. <input type="checkbox"/> J'ai créé l'ébouli des valeurs propres. <input type="checkbox"/> Je suis capable de justifier le nombre de dimensions retenues. <input type="checkbox"/> J'ai créé un cercle des corrélations. <input type="checkbox"/> J'ai réalisé une projection des individus. <input type="checkbox"/> Je me sens capable d'expliquer les caractéristiques des individus et leurs différences.	
	Le support de présentation	<input type="checkbox"/> J'ai respecté le déroulé demandé par Patrick (contexte, démarche, résultats). <input type="checkbox"/> Je suis satisfait de la pertinence des recommandations de cluster(s)/pays pour l'entreprise. <input type="checkbox"/> Je suis satisfait de la structure de ma présentation. <input type="checkbox"/> Je vais utiliser des termes compréhensibles pour un public non technique. <input type="checkbox"/> Je suis capable de justifier mes résultats. <input type="checkbox"/> Je me sens capable de justifier mes choix méthodologiques et de répondre à des questions portant sur : <div> <input type="checkbox"/> le choix du nombre de variables synthétiques issues de l'ACP le cercle des corrélations. </div>	

		<ul style="list-style-type: none"><input type="checkbox"/> La différence de prévision entre la CAH et le K-Means, comment l'expliquer ?<input type="checkbox"/> Faut-il utiliser les données réelles ou des données de l'ACP pour faire le clustering ?<input type="checkbox"/> Quel est le nombre optimal de clusters ?	
--	--	--	--