

## Module : analyser des données avec Python

### Brief projet n°1

Vous êtes consultant chez CarRent , un concessionnaire de voiture multimarque spécialisé dans la vente de voiture d'occasion. Votre manager souhaite augmenter son CA l'année prochaine en investissant dans l'achat de voiture d'occasion les plus tendances du marché. Pour ce faire il a besoin de comprendre sur quels critères un client va se décider à acheter : est ce la marque, la couleur , le modèle ect... ? Votre mission sera donc lui proposer une analyse détaillée de ses ventes de l'année en cours. A l'issu de cette analyse vous devrez lui proposer des recommandations claires , et apporter des informations précises sur les goûts et critères de choix de ses clients.

#### Partie 1 : nettoyage de données

Le service informatique vous livre les données dont il dispose pour l'année en cours :

[donnees\\_CarRent](#) mais vous vous rendez vite compte que celui-ci n'est pas exploitable tel quel et qu'il va falloir le nettoyer !

Vous allez donc devoir :

- Supprimer les colonnes non utiles
  - Renommer les colonnes en français pour plus de clarté
  - Les dates sont toutes erronées : l'administrateur vous signale qu'elles sont toutes décalées de 8 années (les dates de ventes comme les dates d'ancienneté de la voiture)
  - Votre manager vous demande de vous concentrer uniquement sur les voitures ayant moins de 10 ans d'ancienneté : vous allez donc devoir supprimer les lignes non utiles
  - Isoler les lignes dans lesquelles il y a des données manquantes et les exporter dans un CSV à part, pour une analyse ultérieure .
  - Créer des classes pertinentes pour le kilométrage des voitures
  - De même que pour les couleurs, il serait judicieux de réduire le nombre de classes car vous avez trop de variantes de couleur différentes
  - Formater la date de vente : vous n'avez pas réellement besoin de savoir autant de détail, la date de vente est suffisante, l'horaire n'est pas nécessaire.
- ➔ A la fin du nettoyage : n'oublier de préciser combien de données vous avez enlevé, et quelle proportion (en %) de votre dataset il vous reste. Et n'oubliez pas de joindre un rapport d'erreur (dans un doc annexe) Est-ce suffisant pour réaliser une analyse ?
- ➔ Vous manque-t-il des informations ?

## **Partie 2 : analyse et visualisation**

Votre jeu de données est enfin nettoyé, vous pouvez donc commencer votre analyse :

- Réaliser pour commencer, une analyse descriptive univariée de votre jeu de données
- Appliquer la méthode des 5 W pour cadrer votre projet et cibler les attentes exactes de votre manager
- Réalisez au moins 2 maquettes différentes à présenter et à valider avec votre manager avant de vous lancer dans la réalisation concrète de votre rapport et des visualisations
- Vous pouvez donc passer à l'analyse multivariée de votre jeu de données pour tenter de répondre le plus précisément aux attentes qui vous sont demandées.

## **Partie 3 : restitution**

- Vous devrez présenter sous forme d'une présentation powerpoint vos conclusions à l'ensemble du personnel CarRent : il y aura donc votre manager mais également l'équipe marketing et commercial , préparez donc bien votre argumentation et votre discours! Vous n'aurez que 15 min pour faire votre présentation