

Projet

L'objectif principal de ce projet consiste, dans un premier temps, à analyser les données de ventes mensuelles d'une entreprise. Ensuite, entreprendre la segmentation des clients, et enfin, nous développer un modèle de machine learning afin de prédire la valeur à vie de chaque client..

Partie A :

La direction marketing vous soumet un dossier nommé **données_ventes** qui contient plusieurs fichiers. Leur souhait est d'avoir des informations pertinentes de leur activités basées sur les données.

1. Avec un script python, récupérer tous les fichiers qui sont dans le dossier et stocker les dans un fichier csv que vous nommez **data** .
2. Charger le fichier csv enregistré et essayer de répondre aux questions suivantes en faisant les transformations nécessaires. Pour des raisons d'analyse, vous pouvez supprimer toutes les valeurs manquantes. Choisissez un type de graphe pour visualiser chaque résultat.
3. Quel est le mois durant le quel l'entreprise a réalisé le meilleur chiffre d'affaire ?
4. Dans quelle ville, l'entreprise a enregistré le maximum de commandes ?
5. En quel moment doit on faire une campagne de publicité pour avoir plus de ventes ?
6. Quel est le produit qui est plus vendu ? Essayer d'expliquer pourquoi il est plus vendu ?
7. Quelles sont les combinaisons de produits qui se vendent le plus ?
8. Vous pourriez répondre à d'autres questions si nécessaires.

Partie B :

Le département marketing souhaite lancer une campagne de marketing. Ils souhaitent identifier les segments de clientèle les plus rentables pour cibler les efforts marketing et publicitaires de manière plus précise. Votre rôle consiste à diviser les clients de la société en groupes distincts en fonction de certaines caractéristiques, telles que les préférences, les comportements d'achat, ...

1. Charger à nouveau le fichier **data**.
2. Faites le nettoyage et les transformations nécessaires pour faire une segmentation.
3. En utilisant un algorithme de machine learning, faire une segmentation des clients et justifier le nombres de classes obtenues.
4. Décrire pour chaque classe en précisant sa proportion sur le chiffre d'affaires de la société et les produits plus achetés par les clients de chaque classe.

Partie C :

L'équipe marketing très satisfaite de vos travaux précédentes, juge maintenant impératif de mieux comprendre la valeur que chacun des clients apporte à long terme à la société. C'est pourquoi l'équipe marketing souhaite collaborer avec vous les experts en données pour construire un modèle de prédiction de la Valeur à Vie du Client (CLV). Pour cela, elle fournit un autre jeu de données (**customer_segmentation.csv** plus volumineux pour entrainer votre modèle.

1. Charger à nouveau le fichier **customer_segmentation.csv**.
2. Faire les transformations nécessaires pour avoir des données rangées dans l'ordre chronologique et vous allez prédire la valeur à vie d'un client dans **Trois mois**.
3. Appliquer plusieurs modèles de regression pour faire la prédiction et comparer les résultats.

4. Le meilleur modèle obtenu à la dernière question sera enregistré et déployé dans une application web en utilisant **Flask** ou **Streamlit**. Cette application pourra aider les décideurs à utiliser votre modèle mise en place.