
1^{ère} Année Master Science de données et Aide à la décision (S1)

Module: Python pour les sciences de données

Projet2 (Analyse des données pour un système de prédiction)

L'objectif de ce projet est d'utiliser Python avec les bibliothèques Pandas, Matplotlib, Numpy, Tkinter, Seaborn et Scipy pour effectuer une analyse statistique des données. Le projet est à réaliser sur Spyder ou Jupyter.

Ce projet doit être réalisé par un groupe de 4 étudiants.

Attention au plagiat : si deux groupes présenteront le même projet, un zéro sera attribué aux deux groupes.

1/ Contexte

Nous avons dans le dossier datasets/SalesAnalysis des données de ventes d'une entreprise sur l'année 2019. Vous avez un fichier csv par mois.

Chaque fichier contient les colonnes suivantes:

- OrderID: l'identifiant unique de la commande
- Product: La désignation du produit
- Quantity Ordered: la quantité vendue
- Price each: le prix unitaire
- order date : La date d'achat
- Purchase adress: l'adresse de la boutique de vente

Avant de répondre, il faut afficher le dataset de l'année avec un graphe histogramme.

2/ Fonctionnalités souhaitées (sous forme de tableau et graphe)

Nous allons analyser ces données pour répondre aux questions suivantes :

- Quel est le meilleur mois de vente ? et Quel est le chiffre d'affaires de ce mois ?
- Dans quelle ville a-t-on le plus vendu de produits ? et le moins vendu ? afficher les résultats avec des cartes géographiques.
- A quelle heure devons-nous passer de la publicité pour augmenter nos ventes ? quelles sont les heures à éviter ?
- Quels sont les produits qui sont souvent achetés ensemble ? et les produits les moins achetés ?
- Quel produit a-t-on le plus vendu et pourquoi ?
- Quel est le profit de l'entreprise durant l'année ?
- Quel sont les produits à vendre le mois de janvier 2020 (prédiction : utilisation d'un algorithme de machine Learning)

3/ Interface du projet

L'utilisation des interfaces graphiques est obligatoire, votre projet affiche en premier une fenêtre avec un bouton pour charger les datasets, et chaque question correspond à un bouton, un clic sur le bouton une autre fenêtre affiche le résultat.

4/ Datasets :

Les jeux de données sont disponibles sur la plateforme

5/ Consignes

Les étudiants doivent présenter leur projet d'une manière pédagogique en fournissant les explications nécessaires (utilisation des bibliothèques, appel aux méthodes, établissement des diagrammes,etc.) , ces explications doivent précéder chaque étape du projet.

Remise : Début janvier 2024