

Construction d'un Tableau de Bord E-commerce avec Recommandations

Contexte du Challenge

Ce challenge vise à permettre à l'équipe IA-DATA de développer un tableau de bord interactif pour une plateforme e-commerce. Ce tableau permettra aux gestionnaires de suivre les performances commerciales, d'analyser les ventes, de comprendre les comportements des clients, tout en fournissant des recommandations pour optimiser les opérations. Les participants auront une semaine pour finaliser le projet.

Objectif Général

Construire un tableau de bord e-commerce capable de :

- Suivre les performances commerciales et les niveaux de stock.
- Appliquer des techniques de segmentation client pour une meilleure compréhension du comportement d'achat.
- Générer des prédictions de ventes et des recommandations pour l'optimisation des opérations.

Détails du Challenge

- **Durée** : 6 Jours
- **Date de Rendu** : Lundi 23/09/24 matin à 7h12
- **Technologies Utilisées** : Python (Pandas, NumPy, Matplotlib), Scikit-learn, Machine Learning, Tableau croisé dynamique, Régression linéaire

Planification Hebdomadaire

Jour 1 : Introduction et Exploration des Données

- **Objectif 1 (Débutant)** : Utiliser Pandas pour charger et manipuler les données.
 - *Challenge* : Charger et afficher les premières lignes d'un fichier CSV de ventes (produits, clients, prix, quantités).
- **Objectif 2 (Moyen)** : Nettoyer les données pour l'analyse.
 - *Challenge* : Supprimer les valeurs manquantes et formater les dates et prix.
- **Objectif 3 (Difficile)** : Ajouter des fonctionnalités aux données.
 - *Challenge* : Créer des colonnes comme "Chiffre d'affaires par produit" (prix * quantité vendue).

Analyse Exploratoire des Données

- **Objectif 4 (Débutant)** : Visualiser les données avec Pandas et Matplotlib.
 - *Challenge* : Graphique des ventes mensuelles et total des ventes par mois.
- **Objectif 5 (Moyen)** : Analyse approfondie avec des tableaux croisés dynamiques.

- *Challenge* : Tableau croisé des ventes par catégorie de produit.
- **Objectif 6 (Difficile)** : Identifier les tendances.
 - *Challenge* : Visualiser les produits les plus vendus et leur impact sur le chiffre d'affaires.

Jour 2 : Calculs Statistiques et KPI

- **Objectif 1 (Débutant)** : Utiliser NumPy pour des calculs statistiques simples.
 - *Challenge* : Calculer la moyenne, la médiane, l'écart-type des prix.
- **Objectif 2 (Moyen)** : Mettre en place des indicateurs de performance (KPI).
 - *Challenge* : Suivre les KPIs tels que le total des ventes, le profit net, et afficher sur le tableau de bord.
- **Objectif 3 (Difficile)** : Suivi des performances des produits.
 - *Challenge* : Créer des alertes pour les produits avec des ventes faibles.

Jour 3 : Segmentation Client et Analyse des Comportements d'Achat

- **Objectif 1 (Débutant)** : Segmenter les clients par fréquence d'achats.
 - *Challenge* : Créer des groupes (clients réguliers, occasionnels, nouveaux).
- **Objectif 2 (Moyen)** : Analyser le comportement par segment.
 - *Challenge* : Tableau croisé pour voir les catégories de produits les plus achetées par segment.
- **Objectif 3 (Difficile)** : Implémenter une analyse RFM (Récence, Fréquence, Montant).
 - *Challenge* : Classer les clients et afficher les résultats sous forme de tableau.

Jour 4 : Prédiction des Ventes avec Régression Linéaire

- **Objectif 1 (Débutant)** : Introduction à Scikit-learn et régression linéaire.
 - *Challenge* : Implémenter une régression simple pour prédire les ventes futures.
- **Objectif 2 (Moyen)** : Optimiser les prédictions avec plusieurs variables.
 - *Challenge* : Utiliser la régression multiple (catégories de produits, saisonnalité).
- **Objectif 3 (Difficile)** : Améliorer la précision.
 - *Challenge* : Validation croisée et ajustement des hyperparamètres.

Jour 5 : Optimisation des Stocks et des Prix

- **Objectif 1 (Débutant)** : Analyser les fluctuations de stock.
 - *Challenge* : Identifier les produits en surstock ou rupture et visualiser dans le tableau de bord.
- **Objectif 2 (Moyen)** : Prédire les besoins futurs en stock.

- *Challenge* : Modèle de régression pour prédire les besoins en stock.
- **Objectif 3 (Difficile)** : Stratégie de prix dynamique.
 - *Challenge* : Utiliser un algorithme pour recommander des prix optimisés.

Jour 6 : Intégration et Finalisation du Tableau de Bord

- **Objectif 1 (Débutant)** : Créer un tableau de bord interactif avec Pandas et Matplotlib.
 - *Challenge* : Afficher les ventes mensuelles, le nombre de clients, produits les plus vendus.
- **Objectif 2 (Moyen)** : Ajouter des filtres interactifs.
 - *Challenge* : Créer un système de filtrage pour les critères (catégorie produit, segment clients).
- **Objectif 3 (Difficile)** : Intégrer des recommandations en temps réel.
 - *Challenge* : Automatiser des recommandations et les afficher dans le tableau de bord.

Livrables

À la fin du challenge, l'équipe devra livrer un tableau de bord interactif contenant :

1. Une analyse des performances commerciales.
2. Une segmentation des clients et leur comportement d'achat.
3. Des prédictions et recommandations pour optimiser les opérations commerciales.

Ce projet permettra à l'équipe d'acquérir des compétences en manipulation de données, machine learning et visualisation tout en apportant des solutions pratiques pour une plateforme e-commerce.

Deadline

- **Date de Rendu** : Lundi matin à 7h

Ressources :

Dataset : <https://github.com/TechSeed-Academy/Ressources.git> branch ai_data

Ressources Générales

1. **Documentation Python** : <https://docs.python.org/3/>
2. **Jupyter Notebook** (pour un environnement interactif) : <https://jupyter.org/>

Introduction et Exploration des Données

- **Pandas (Manipulation des données)** :
 - Documentation Pandas
 - Tutoriel : Pandas 101

- **Nettoyage des données :**
 - Tutoriel : Cleaning Data with Pandas

Analyse Exploratoire des Données

- **Matplotlib (Visualisation) :**
 - Documentation Matplotlib
 - Tutoriel : Matplotlib Basics
- **Tableaux Croisés Dynamiques avec Pandas :**
 - Pivot Table Guide

Calculs Statistiques et KPI

- **NumPy (Calculs statistiques) :**
 - Documentation NumPy
 - Tutoriel : Introduction to NumPy
- **KPI :**
 - Tutoriel : How to Build a KPI Dashboard

Segmentation Client et Analyse des Comportements d'Achat

- **Segmentation Client avec Pandas :**
 - Tutoriel : Customer Segmentation with Pandas
- **Analyse RFM (Récence, Fréquence, Montant) :**
 - Tutoriel : RFM Analysis

Prédiction des Ventes avec Régression Linéaire

- **Scikit-learn (Régression Linéaire) :**
 - Documentation Scikit-learn
 - Tutoriel : Linear Regression with Scikit-learn
- **Validation croisée :**
 - Tutoriel : Cross-Validation with Scikit-learn

Optimisation des Stocks et des Prix

- **Optimisation des stocks et régression :**
 - Tutoriel : Supply Chain Optimization
- **Stratégie de prix dynamique :**
 - Article : Dynamic Pricing with Machine Learning

Intégration et Finalisation du Tableau de Bord

- **Création de Tableau de Bord Interactif :**

- Tutoriel : Building Interactive Dashboards with Plotly Dash
- Tutoriel : Interactive Dashboards with Matplotlib and Pandas

Outils Complémentaires

- **Google Colab** (alternative à Jupyter, cloud-based) : <https://colab.research.google.com/>
- streamlit
- **GitHub** (pour gestion de version) : <https://github.com/>