

Séance 1 B3YNOV :

Prise de Décision Basée sur les Données

Analyse et Prédiction des Prix des Diamants

Durée estimée : 4-8 heures

Contexte

Vous êtes chargés d'une mission critique pour l'entreprise : poser les bases d'un processus de prise de décision stratégique. Dans cette phase initiale, nous nous concentrons sur l'analyse des données et sur la création d'un modèle prédictif basique qui nous aidera à mieux comprendre les dynamiques influençant le prix des diamants. Ce travail préliminaire orientera les prochaines étapes du projet, prévues dans quelques mois, où un modèle plus sophistiqué sera développé.

Votre objectif principal est de fournir des analyses approfondies et des recommandations basées sur les données pour soutenir les décisions stratégiques.

Objectifs principaux

- Explorer et analyser les données pour identifier les tendances et relations clés.
- Développer un modèle de prédiction basique pour guider les prises de décisions initiales.
- Fournir des insights exploitables pour orienter les futures étapes du projet.

Mission : Message de M. Jean, Directeur du Projet

Bonjour à tous et à toutes, dans cette première phase, il est primordial de se concentrer sur l'analyse des données et sur la prise de décisions éclairées. Bien que le modèle de prédiction que vous développerez aujourd'hui soit basique, il jouera un rôle clé en fournissant des indications pour le développement d'un modèle plus avancé dans quelques mois.

Points importants :

- Priorité aux analyses :**
 - Plus que la complexité du modèle, **c'est la clarté et la pertinence des analyses qui comptent à ce stade**. Vous devez bien comprendre les relations entre les caractéristiques des diamants et leur prix.
 - Expliquez vos choix méthodologiques et mettez en avant des recommandations concrètes issues de vos analyses.
- Modèle basique :**
 - Un modèle simple (ex. : régression linéaire) suffit pour cette phase.
 - L'objectif est de fournir des insights pour guider les prochaines étapes du projet, pas de livrer un produit final.
- Prise de décision basée sur les données :**
 - Identifiez les facteurs clés influençant le prix des diamants.
 - Formulez des recommandations basées sur vos observations pour optimiser nos processus.
- Sécurité des données et collaboration :**
 - Respectez les consignes sur la confidentialité des données.
 - Utilisez les outils mis à votre disposition (ChatGPT, Gemini) juste pour maximiser votre efficacité.

Ce que j'attends de vous :

- Une analyse approfondie des données pour identifier les principales tendances.
- Un modèle basique avec une interprétation claire des résultats.
- Des recommandations exploitables pour orienter les décisions stratégiques de l'entreprise.

Bon courage et soyez méthodiques, Jean

Plan approximatif de la Séance :

- 1. Introduction**
 - Présentation des objectifs de la séance.
 - Mise en avant de l'importance des analyses pour guider les décisions.
- 2. Exploration des données**
 - Charger et examiner la base de données diamonds.csv.
 - Identifier les tendances générales et les relations entre variables.
- 3. Analyse approfondie**
 - Étudier les relations entre les caractéristiques clés (carat, coupe, couleur, clarté, etc.) et les prix.
 - Produire des visualisations pour appuyer vos observations.
- 4. Modélisation basique**
 - Entraîner un modèle simple (ex. : régression linéaire).
 - Évaluer la performance du modèle avec des métriques basiques (ex. : MAE, RMSE).
- 5. Recommandations et prise de décision**
 - Synthétiser vos analyses en insights exploitables.
 - Proposer des recommandations stratégiques basées sur vos résultats.
- 6. Conclusion**
 - Discussion sur l'utilité des analyses pour orienter la prochaine phase du projet.
 - Réflexion sur les améliorations possibles pour les étapes futures.

Résultat final de la séance :

Une analyse détaillée des données et un modèle basique qui serviront de base pour orienter les décisions stratégiques de l'entreprise et préparer le développement d'un modèle avancé dans le futur. Votre travail aujourd'hui doit insister sur la prise de décision éclairée pour maximiser la valeur ajoutée des données disponibles.