

A decorative graphic on the left side of the slide, consisting of white lines and circles on a blue gradient background, resembling a circuit board or data flow diagram.

INTRODUCTION

BIENVENUE DANS MA PRÉSENTATION FINALE SUR LE PARCOURS QUE J'AI SUIVI EN SCIENCE DES DONNÉES AVEC IBM. AU COURS DE CETTE FORMATION, J'AI EXPLORÉ DIVERS ASPECTS DE L'ANALYSE DE DONNÉES, DES BASES THÉORIQUES AUX APPLICATIONS PRATIQUES. MON OBJECTIF PRINCIPAL ÉTAIT DE DÉVELOPPER DES COMPÉTENCES SOLIDES EN MANIPULATION DE DONNÉES, EN ANALYSE EXPLORATOIRE, ET EN MODÉLISATION PRÉDICTIVE, TOUT EN UTILISANT DES OUTILS AVANCÉS COMME PYTHON ET IBM WATSON STUDIO.

CONTEXTE ET OBJECTIFS

- **Contexte :** Mon voyage a débuté avec une introduction aux fondamentaux de la science des données. L'objectif était de comprendre comment transformer des données brutes en insights précieux, en utilisant des techniques statistiques et d'apprentissage automatique.
- **Objectifs :**
 - Maîtriser les techniques de nettoyage et de préparation des données.
 - Explorer les données à l'aide de visualisations interactives.
 - Construire et évaluer des modèles prédictifs pour des problèmes complexes.

PROJETS ET MÉTHODOLOGIES

- **Projet 1 : Nettoyage et Préparation des Données**

Objectif : Améliorer la qualité des données pour l'analyse.

Méthodologie : Utilisation de Python et de Pandas pour le nettoyage des données, gestion des valeurs manquantes, et correction des valeurs aberrantes.

Résultat : Un jeu de données propres et prêt pour l'analyse exploratoire.

- **Projet 2 : Analyse Exploratoire des Données (EDA)**

Objectif : Découvrir des tendances et des relations dans les données.

Méthodologie : Visualisations avec Matplotlib et Seaborn, exploration interactive avec Plotly.

Résultat : Identification de corrélations significatives et de détection de motifs cachés dans les données.

- **Projet 3 : Modélisation Prédictive**

Objectif : Construire des modèles pour prédire des résultats spécifiques.

Méthodologie : Application de régression logistique et d'arbres de décision avec Scikit-learn, évaluation des modèles avec des métriques comme la précision et le rappel.

Résultat : Modèles prédictifs efficaces, avec des performances démontrées par des courbes ROC et des matrices de confusion.

VISUALISATION ET OUTILS

- **Visualisations Clés :**

Histogrammes et Diagrammes en Boîte : Pour comprendre la distribution des variables.

Diagrammes de Dispersion : Pour explorer les relations entre différentes variables.

Cartes Interactives avec Folium : Pour visualiser les données géographiques de manière dynamique.

- **Outils Utilisés :**

Python : Pour la manipulation et l'analyse des données.

- **IBM Watson Studio :** Pour l'intégration de l'analyse de données et de la modélisation prédictive.

- **Plotly et Folium :** Pour des visualisations interactives et cartographiques.

APERÇUS ET CONCLUSIONS

- **Points importants :**

Corrélations détectées : Découverte de relations entre les variables, par exemple, l'impact des habitudes alimentaires sur la santé.

Modèles Prédicatifs : Réussite de la prédiction du risque de maladies avec une précision remarquable.

- **Conclusions :**

Les données bien nettoyées et préparées sont essentielles pour une analyse efficace.

Les visualisations interactives facilitent la compréhension et la prise de décisions basées sur les données.

Les modèles prédictifs construits sont robustes et offrent des informations précieuses pour des décisions stratégiques.

COMPÉTENCES ACQUISES

- **Techniques :**

Nettoyage et Préparation des Données : Expertise dans l'utilisation de Pandas et Numpy.

Exploration et Visualisation : Maîtrise de Matplotlib, Seaborn et Plotly pour des visualisations interactives.

Modélisation : Compétences en construction de modèles avec Scikit-learn, y compris la régression, les arbres de décision, et les forêts aléatoires.

- **Compétences Transversales :**

Communication des Résultats : Capacité à présenter des insights complexes de manière claire et engageante.

Collaboratif et Innovant : Utilisation de méthodes créatives pour résoudre des problèmes de données complexes.

PROCHAINES ÉTAPES

- **Développement Futur :**

Approfondissement des Connaissances : Exploration de l'apprentissage profond et de l'intelligence artificielle avancée.

Application Pratique : Mise en œuvre de projets réels pour résoudre des problèmes d'entreprise concrets.

- **Points de vue :**

Je suis enthousiaste à l'idée de continuer à développer mes compétences en science des données et de contribuer à des projets innovants qui utilisent les données pour transformer le monde.