

Prédiction des Annulations de Réservations

Réduire les pertes et améliorer l'expérience client



Projet réalisé par : **Bigaëlle GUERRIER**

Contexte

- L'entreprise subit des **pertes financières** dues aux réservations annulées ou incomplètes.
- Objectif : **anticiper les annulations pour réduire les pertes et améliorer l'expérience utilisateur.**
- Parties prenantes : direction, équipes opérationnelles, clients et chauffeurs.

Dataset

- Plus de **150 000 réservations** analysées.
- Variables principales :
 - Type de véhicule
 - Distance du trajet
 - Montant de la course
 - Notes clients et chauffeurs
 - Méthode de paiement
- Valeurs manquantes présentes → nettoyées avant modélisation.

Préparation des données

- Transformation des variables catégorielles en valeurs numériques (encodage).
- Création de la **cible binaire** :
 - Completed = 1
 - Cancelled / Incomplete / No Driver Found = 0

Séparation des données en **train/test** pour évaluer les modèles.

Analyse exploratoire (EDA)

- Visualisation de la répartition des réservations complétées et annulées.
- Corrélations importantes : distance \leftrightarrow montant, notes \leftrightarrow annulations.
- Patterns observés :
 - Certains types de véhicules ont plus d'annulations
 - Les jours et heures influencent le risque
 - Méthodes de paiement plus ou moins fiables

Modèles utilisés

- **Forêt Aléatoire (Random Forest)** : robuste, interprétation simple.
- **XGBoost** : performant pour les relations complexes.
- Évaluation avec : **Accuracy, Precision, Recall, F1-score, AUC**.

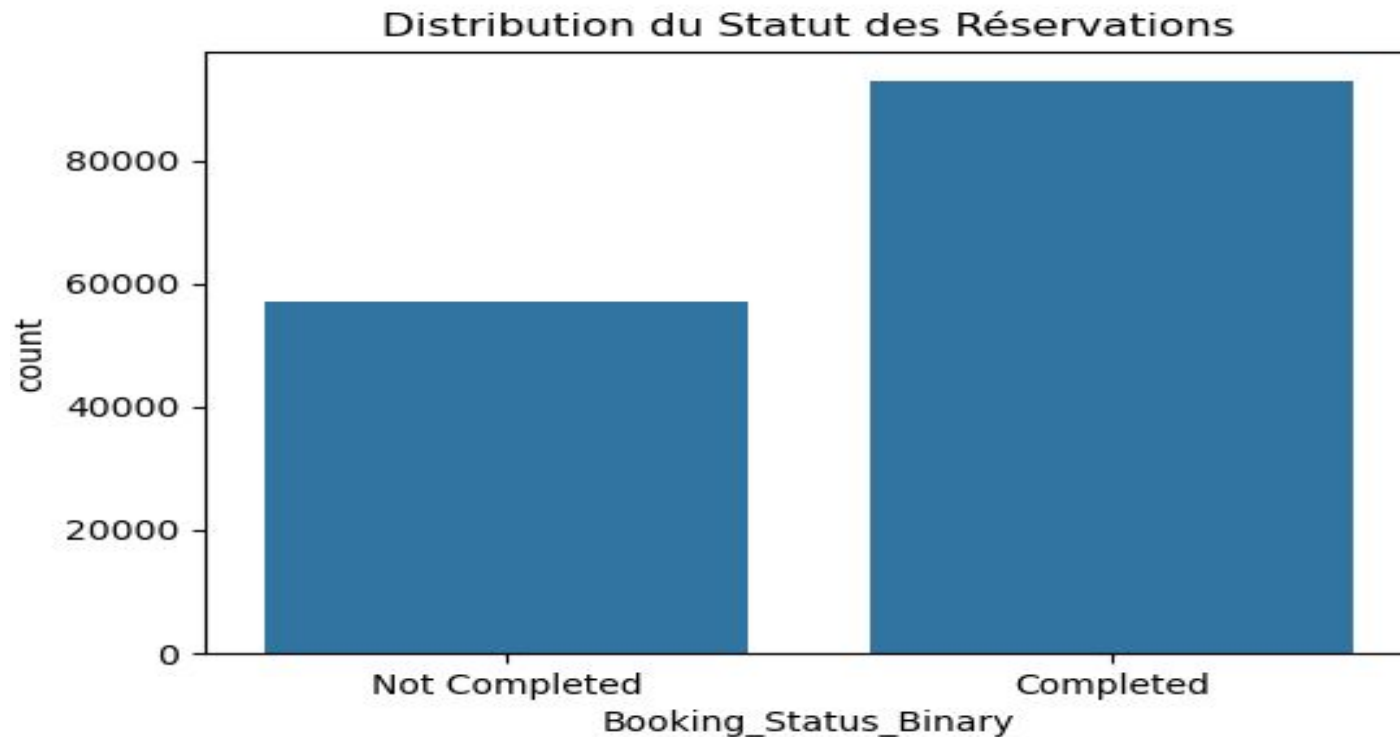
Interprétation avec SHAP

- Permet de **comprendre quelles variables influencent le plus le risque d'annulation**.
- Les 6 variables clés identifiées :
 1. Payment Method
 2. Rides incomplètes
 3. Ride Distance
 4. Booking Value
 5. avg ctat (Customer Rating)
 6. avg vtat (Driver Rating)

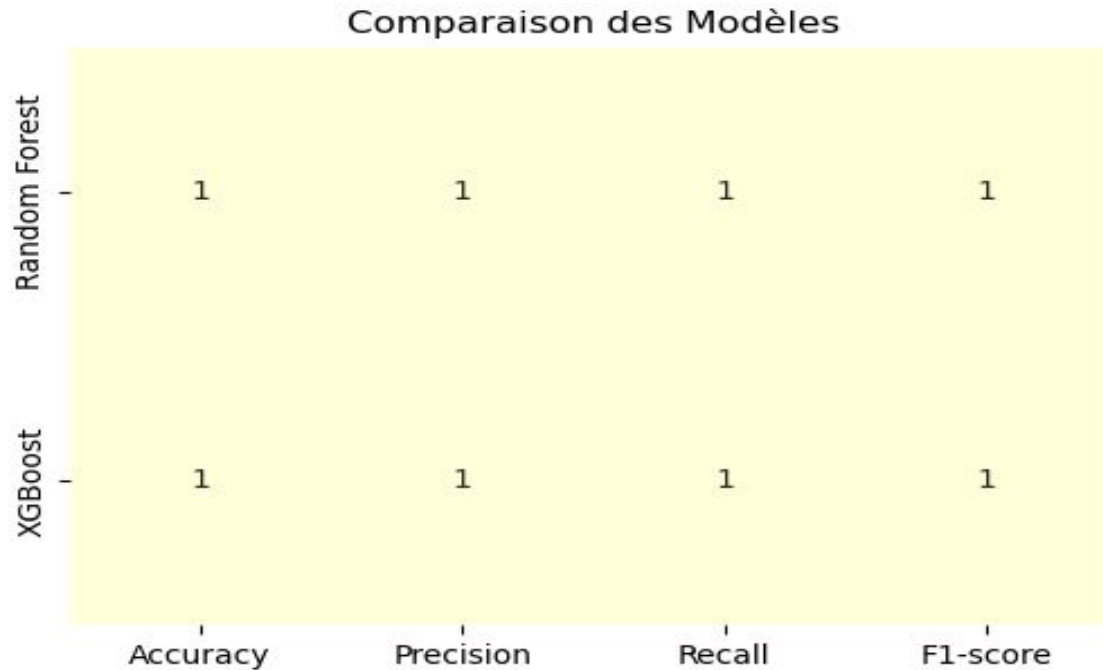
Recommandations business

- Optimiser la méthode de paiement
- Réduire les courses incomplètes
- Gestion des trajets
- Focus sur la valeur de la réservation
- Suivi des notes client et chauffeur

Distribution du Statut des Réservations



Comparaison des Modèles



Conclusions

- Le modèle **prévoit efficacement les annulations**.
- Les features clés offrent des **actions concrètes pour le business**.
- XGBoost légèrement plus performant que la Forêt Aléatoire.

Recommandations business

1. Optimiser les méthodes de paiement.
2. Ajuster la disponibilité des chauffeurs selon les jours à risque.
3. Surveiller les trajets longs et coûteux.
4. Suivi des notes clients et chauffeurs pour réduire les annulations.
5. Mise à jour régulière du modèle avec les nouvelles données.