

Prévision des Revenus de Ventes de Smartphones

Rapport :

Réalisé par :

BENTBIB

Zineb

Sommaire :

- 1. Introduction**
- 2. Environnement Technique :**
- 3. Compétences techniques :**
- 4. Description du travail réalisé :**
 - **Exploitation du jeu de données**
 - **Pré-traitement des données**
 - **Nettoyage de Données**
 - **Analyse Descriptive**
 - **Recommandations business basées sur les insights**
 - **Modèle de prévision journalier des revenus pour chaque smartphone pour la période du 1er janvier 2025 au 31 Mars 2025**
 - **Justifications du choix du modèle**
 - **Les prévisions et leur interprétations**
 - **Méthode d'Evaluation du modèle**
 - **Recommandations opérationnelles**

Introduction :

Dans un monde où les entreprises cherchent à maximiser leurs revenus tout en optimisant leurs ressources, l'analyse prédictive devient un outil indispensable pour la prise de décision. Ce projet s'inscrit dans cette dynamique en mettant en œuvre des techniques avancées de science des données pour prédire les revenus journaliers de différents modèles de smartphones.

L'objectif principal de ce projet est de développer un modèle robuste, capable d'exploiter des données historiques et des variables exogènes (marketing, pouvoir d'achat, événements technologiques, etc.), afin d'offrir des prévisions fiables pour chaque modèle de smartphone et chaque région. Ces prédictions permettent non seulement d'anticiper les tendances de vente, mais aussi de guider les stratégies opérationnelles, telles que la gestion des stocks, les campagnes marketing et les ressources logistiques.

Environnement Technique :

Pour mener à bien ce projet, plusieurs outils et technologies ont été utilisés, garantissant une analyse rigoureuse des données et la mise en œuvre d'un modèle prédictif robuste. Voici les ressources nécessaires

Compétences techniques :

Python

Bibliothèques :

Pandas : Manipulation et analyse des données tabulaires, y compris le nettoyage et le prétraitement.

NumPy : Calculs numériques rapides, utilisé pour la manipulation des tableaux et des données.

Matplotlib et Seaborn : Génération de visualisations pour explorer les données et interpréter les résultats.

Scikit-learn : Mise en œuvre du modèle de machine learning (Random Forest Regressor) et évaluation des performances.

Description du travail réalisé :

- **Exploitation du jeu de données :**

Les données historiques de ventes ont été chargées à partir d'un fichier CSV en utilisant la bibliothèque Pandas.

Le dataset fourni contient les ventes journalières de smartphones de 2019 à 2024 pour 8 villes françaises et 3 modèles de smartphones (iPhone Pro, Kaggle Pixel 5, Planet SX). Il inclut des données de revenus ainsi que plusieurs variables exogènes influençant les ventes, telles que les campagnes marketing, la concurrence, la satisfaction client, le pouvoir d'achat local, la météo, les événements technologiques, la maturité du réseau 5G, l'affluence en magasin et l'état des transports publics. Ce dataset permettra d'analyser les tendances de vente et de développer un modèle de prévision des revenus pour le premier trimestre 2025.

- **Pré-traitement des données :**

Le nettoyage des données constitue une étape essentielle dans tout projet de science des données. Avant d'entamer les analyses exploratoires ou la modélisation, il est crucial de s'assurer que les données sont complètes, cohérentes et prêtes à être exploitées.

Dans le cadre de ce projet, les données brutes contiennent des valeurs manquantes, des anomalies potentielles, et des incohérences dans les types de certaines colonnes

- **Nettoyage de Données :**

Le dataset contenait des valeurs manquantes dans des colonnes numériques et catégoriques.

Colonnes numériques : Imputation avec la moyenne (ex. marketing_score, store_traffic, competition_index, customer_satisfaction, purchasing_power_index)

Colonnes catégoriques : Imputation avec la modalité la plus fréquente (ex. weather_condition, tech_event, 5g_phase, public_transport).

Revenus : Suppression des lignes avec des valeurs manquantes dans iPhone_Pro_revenue, Kaggle_Pixel_5_revenue, et Planet_SX_revenue.

Dates : Conversion en type datetime pour permettre des analyse temporelles

Résultats :

- Toutes les valeurs manquantes ont été imputées ou supprimées.
- Aucune valeur manquante restante dans le dataset.

Taille du dataset :

- Nombre de lignes après nettoyage : 14,606 (2,930 lignes supprimées pour des valeurs manquantes critiques).

- **Analyse Descriptive :**

Des statistiques descriptives ont été calculées pour comprendre les distributions des variables clés comme les revenus des smartphones.

Statistique	jPhone Pro Revenue (€)	Kaggle Pixel 5 Revenue (€)	Planet SX Revenue (€)
Count	14,606	14,606	14,606
Mean	83,443.92	40,366.47	21,738.65
Std	56,633.54	26,585.13	14,126.05
Min	0.00	1,576.18	1,066.71
25%	45,411.20	22,264.23	12,270.26
50% (Median)	70,504.80	33,735.42	18,184.04
75%	103,735.32	49,791.18	26,709.87
Max	536,859.73	222,601.22	103,250.71

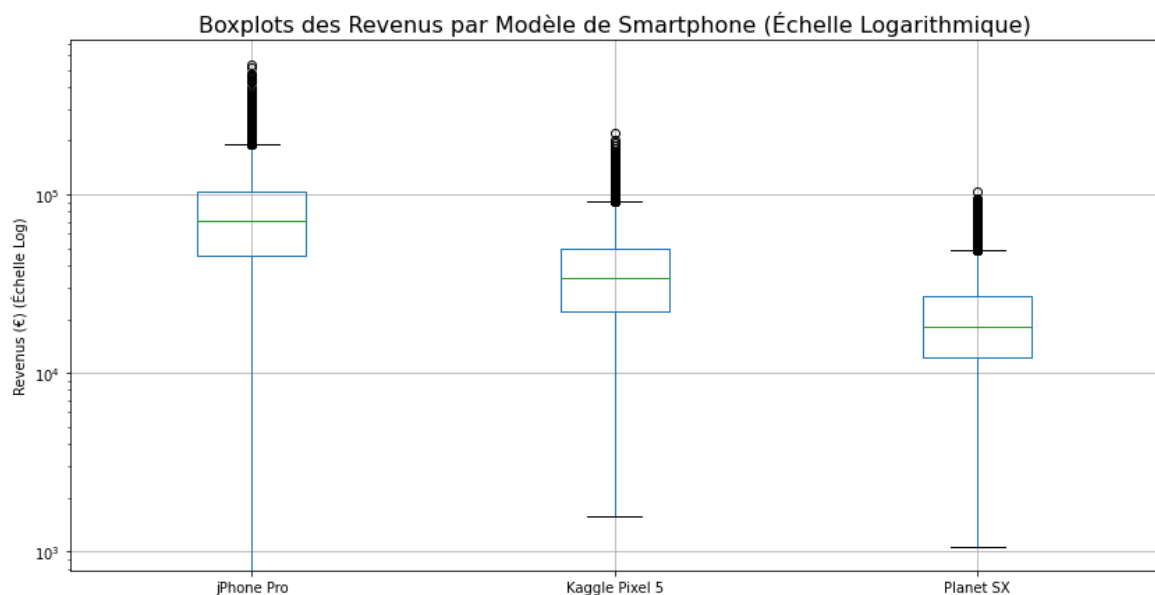
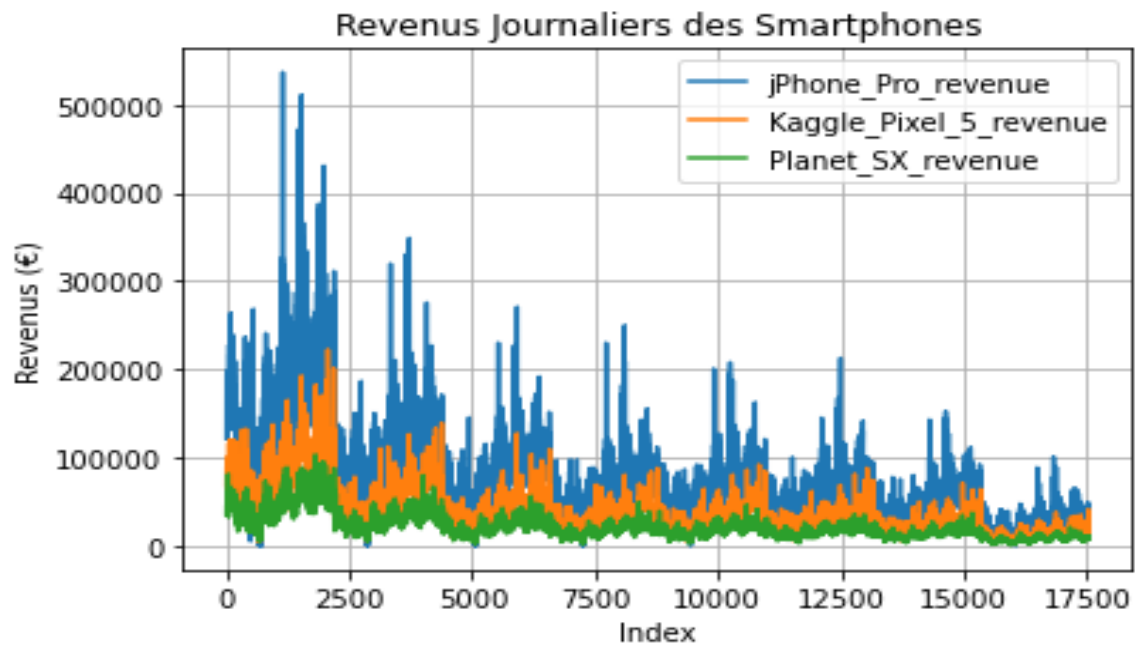
Observations :

- Le iPhone Pro génère les revenus moyens les plus élevés (83,443.92 €), suivi par le Kaggle Pixel 5 (40,366.47 €) et le Planet SX (21,738.65 €).

Analyse globale :

Le iPhone Pro domine en termes de revenus moyens et maximums, mais ses ventes sont plus sujettes à des fluctuations importantes. En revanche, le Kaggle Pixel 5 et le Planet SX présentent des revenus plus stables, bien qu'inférieurs.

Insights : Les revenus montrent des opportunités d'optimisation pour les segments premium et milieu de gamme. Les modèles entrée de gamme pourraient nécessiter des efforts marketing dans des régions spécifiques pour augmenter leurs ventes.



Observations :

jPhone Pro :

- Dispersion élevée :.
- Nombre important d'outliers : Beaucoup de points isolés en haut montrent des pics de revenus (événements spéciaux, promotions, lancements).
- Revenus élevés : La médiane est plus haute que celle des autres modèles, confirmant sa domination sur les revenus.

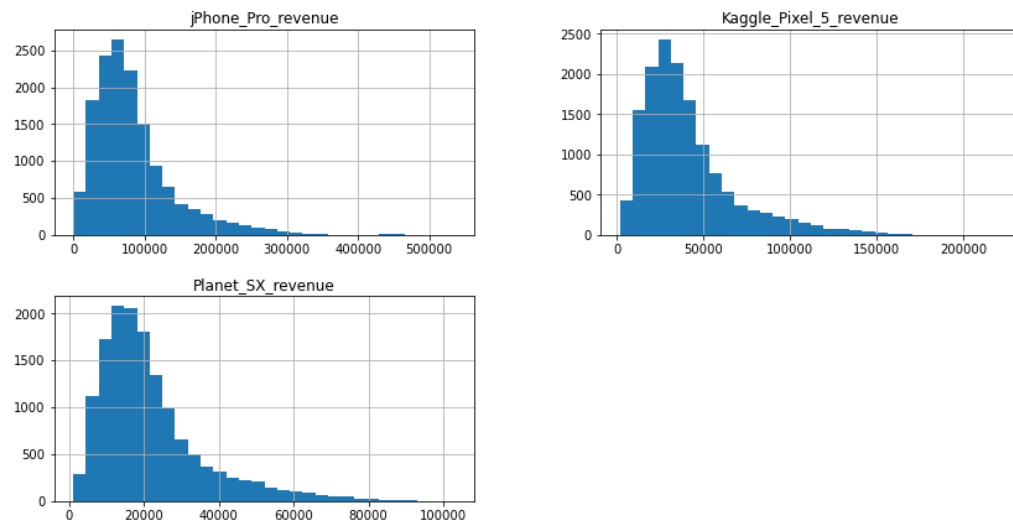
Kaggle Pixel 5 :

- Dispersion modérée : La boîte est plus petite que celle du iPhone Pro, indiquant des revenus relativement stables.
- Outliers présents : Quelques points isolés en haut, mais moins nombreux que pour le iPhone Pro.
- Position médiane intermédiaire : Les revenus du Kaggle Pixel 5 se situent entre ceux du iPhone Pro et du Planet SX.

Planet SX :

- Dispersion la plus faible : La boîte est étroite, montrant que les revenus sont concentrés dans une plage plus restreinte.
- Revenus faibles : La médiane est la plus basse parmi les trois modèles.
- Quelques outliers : Quelques points isolés en haut indiquent des revenus exceptionnellement élevés, bien qu'ils soient rares

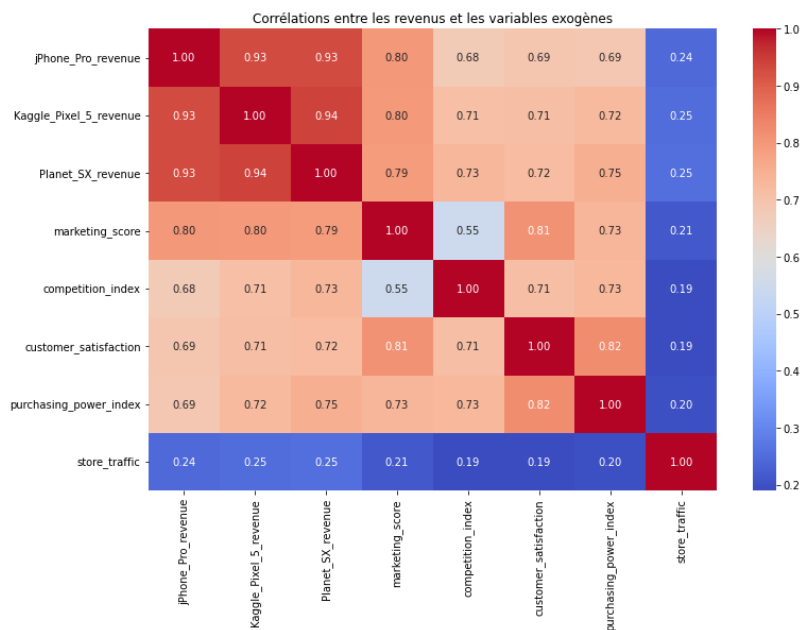
Distributions des Revenus par Modèle de Smartphone



Observations :

- iPhone Pro : Variabilité élevée et pics de revenus
- Kaggle Pixel 5 : Revenus plus stables, moins de fluctuations importantes.
- Planet SX : Revenus faibles et constants, peu de fluctuations.

Calcul des corrélations :



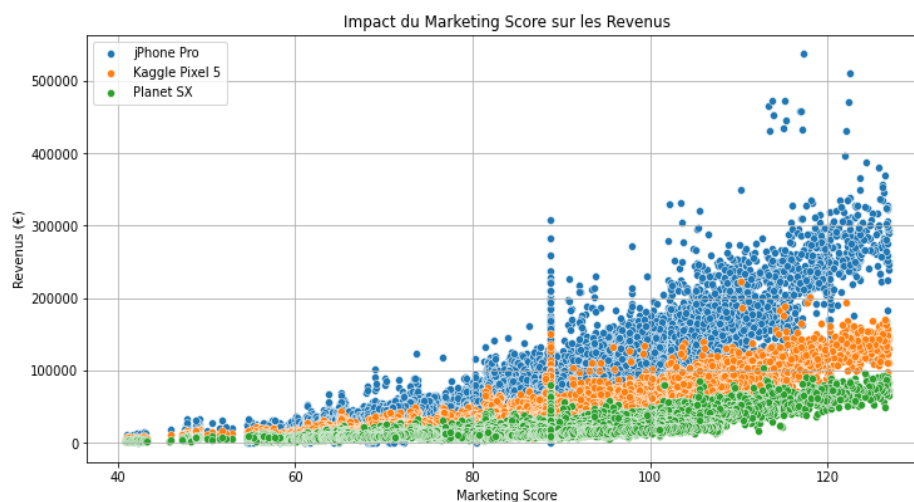
Observations :

- Les revenus des trois modèles de smartphones sont fortement corrélés entre eux

Les revenus des trois modèles sont fortement corrélés (corrélation > 0.90), suggérant des influences globales communes (saisonnalité, promotions).

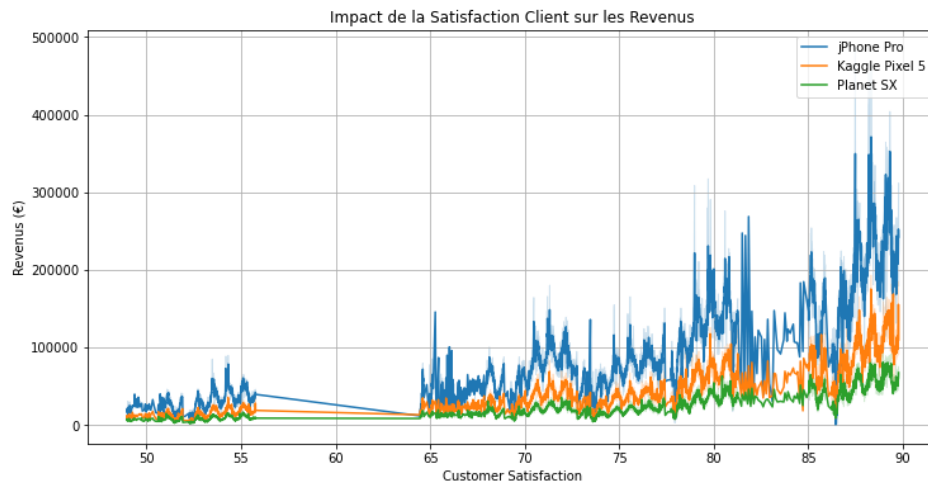
Facteurs influençant les revenus :

- Marketing Score : Influence majeure (corrélation ≈ 0.80). Les campagnes marketing impactent fortement les revenus des trois modèles.
- Satisfaction Client : Corrélation modérée à forte (≈ 0.70). Plus les clients sont satisfaits, plus les revenus augmentent.
- Pouvoir d'Achat : Corrélation significative ($\approx 0.70-0.75$), particulièrement importante pour le modèle Planet SX.
- Concurrence : Corrélation modérée (≈ 0.70), montrant un impact négatif relatif des compétiteurs.



Observations :

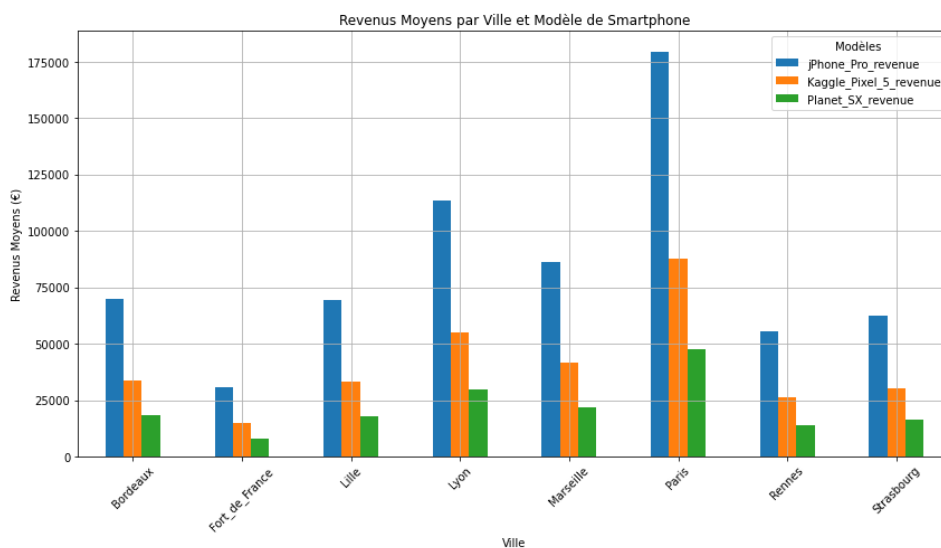
- Les revenus augmentent systématiquement avec le Marketing Score pour les trois modèles de smartphones.
- Cette relation est particulièrement marquée à partir d'un score de 80.



- Observations :

Les revenus augmentent avec l'indice de satisfaction client pour les trois modèles de smartphones.

Cette tendance est particulièrement marquée lorsque l'indice de satisfaction dépasse 70.



- Observations :

Paris affiche les revenus moyens les plus élevés pour les trois modèles de smartphones, indiquant une forte demande dans cette ville. Cela peut être dû à un pouvoir d'achat plus élevé et une plus grande densité de population.

Les villes avec un pouvoir d'achat élevé (comme Paris et Lyon) montrent une préférence plus marquée pour les modèles premium, tandis que les villes avec un pouvoir d'achat plus bas (comme Fort-de-France) génèrent des revenus plus équilibrés entre les segments.

- **Recommandations business basées sur les insights :**

Paris, Lyon, et Marseille génèrent les revenus les plus élevés. Investir dans des campagnes marketing ciblées pour ces villes, surtout pour le modèle premium iPhone Pro.

Dans les villes comme Fort-de-France et Rennes, prioriser les modèles économiques (Planet SX) et lancer des promotions attractives.

Les événements technologiques boostent les revenus, particulièrement pour le iPhone Pro. Planifier des campagnes autour de ces événements

- **Modèle de prévision journalier des revenus pour chaque smartphone pour la période du 1er janvier 2025 au 31 Mars 2025 :**

J'ai opté pour le Random Forest Regressor, un modèle d'apprentissage supervisé basé sur des arbres de décision. Ce modèle est particulièrement efficace pour traiter des ensembles de données complexes avec des interactions non linéaires entre les variables explicatives et les cibles.

.1-Justifications du choix du modèle :

Les variables utilisées incluent des données numériques (marketing_score, purchasing_power_index) et des variables catégoriques (city, weather_condition).

Donc Pourquoi Random Forest ? Il gère facilement ces deux types de données sans nécessiter de transformation complexe.

Le modèle réduit le risque de surapprentissage (overfitting) grâce à l'agrégation de plusieurs arbres.

2. Les prévisions et leur interprétations :

Les prédictions montrent des variations logiques :

iPhone Pro : Génère des revenus élevés, en ligne avec le segment premium.

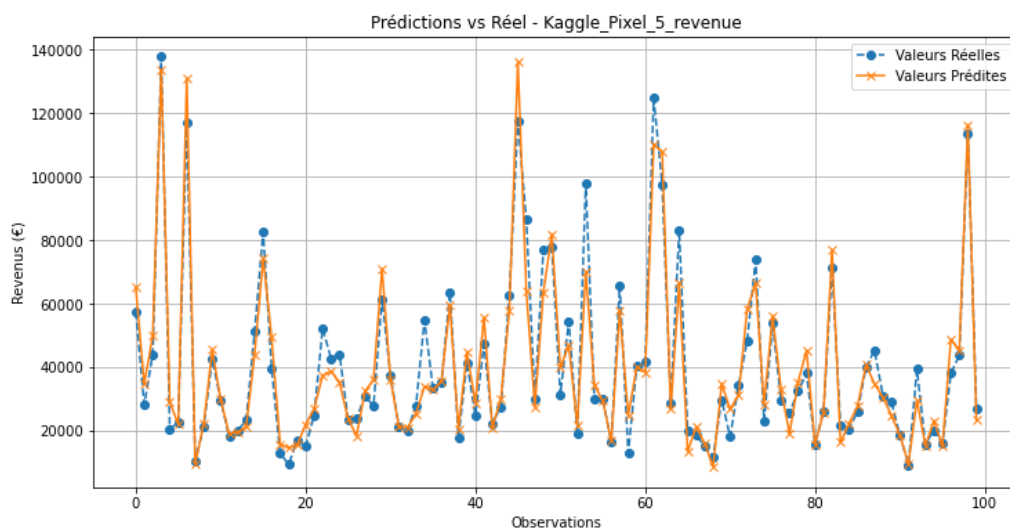
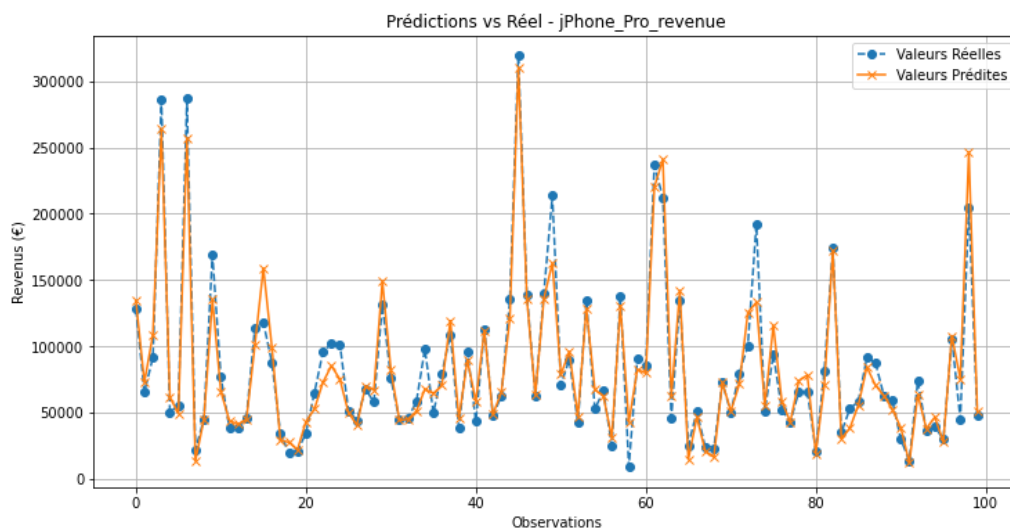
Kaggle Pixel 5 : Positionné au milieu, ses revenus sont cohérents avec un modèle de milieu de gamme.

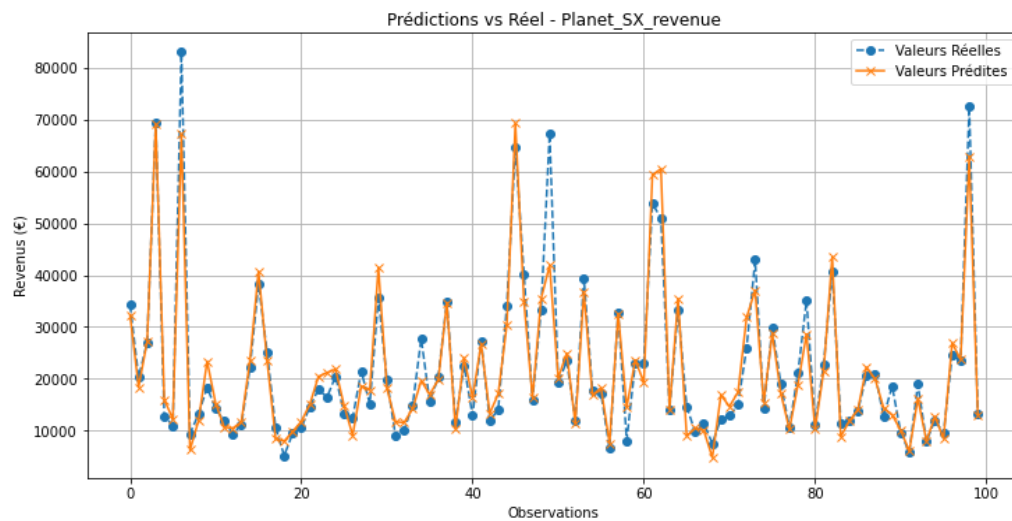
Planet SX : Génère des revenus inférieurs, logique pour un modèle d'entrée de gamme.

3-Méthode d'Evaluation du modèle :

Le modèle de Random Forest offre de bonnes performances globales, particulièrement pour les modèles haut de gamme (iPhone Pro) et milieu de gamme (Kaggle Pixel 5). Toutefois, des ajustements sont nécessaires pour mieux prédire les ventes du Planet SX

En gros , Le modèle est performant et offre une bonne capacité de généralisation.





4-Recommandations opérationnelles :

- Prioriser les stocks des modèles premium dans les zones à fort pouvoir d'achat.
- Concentrer les promotions des modèles économiques dans les régions à faible revenu.
- Renforcer les campagnes marketing autour des événements technologiques pour maximiser les revenus du iPhone Pro.
- intégrer des données supplémentaires sur les préférences locales et les tendances de consommation.