



LABORATOIRE D'ECONOMIE

RAPPORT DU PROJET

---

## Prédiction des ventes d'un produit pharmaceutique

---

**Réalisé par :**

Gaëtan **BLECON**

Modeste E. **SOMAVO**

Reda M. **GAABOURI**

Mathis **VIE**

**Sous la supervision de :**

Ismaël Arnaud **BAGORO**

Mai 2024

# Table des matières

I. Introduction .....	1
II. Revue méthodologique .....	2
2.1. ARMA, ARIMA et SARIMA .....	2
2.2. Lissage exponentiel : Simple, Double, Triple.....	3
2.3. Prophet .....	4
2.4. Réseaux de Neurones Artificiels.....	4
III. Exploration des données .....	7
3.1. Valeurs manquantes.....	7
3.2. Visualisation de la série .....	7
3.3. Analyse de la Tendance et de la Saisonnalité .....	8
3.4. Analyse de la stationnarité .....	11
IV. Application des modèles et résultats .....	14
4.1. SARIMA .....	14
4.2. TES (Triple Exponential Smoothing) .....	15
4.3. Prophet .....	15
4.4. Prévision finale .....	16
V. Conclusion .....	18
Annexe.....	19

# I. Introduction

## CONTEXTE

La prévision des ventes est un élément crucial pour les entreprises, permettant d'anticiper les besoins futurs plutôt que de simplement comprendre les tendances passées. Contrairement aux modèles explicatifs qui cherchent à comprendre les causes sous-jacentes, les modèles de prévision se concentrent sur l'estimation des niveaux futurs de demandes basés sur des données historiques. Ce projet se situe dans ce cadre, en ciblant spécifiquement le secteur pharmaceutique.

## OBJECTIF

Notre objectif est de prédire les ventes d'un produit pharmaceutique pour la période de juillet 2022 à juin 2023. Nous évaluerons plusieurs méthodes de prévision pour déterminer la plus efficace pour nos données spécifiques.

## PLAN

Nous débuterons par une revue méthodologique des différentes approches de prévision, traditionnelles comme modernes. Ensuite, nous explorerons les données de ventes pour en identifier les caractéristiques principales, telles que les tendances et la saisonnalité. Ceci nous permettra d'appliquer et de comparer les méthodes sélectionnées sur un échantillon divisé entre données d'entraînement et de test, pour évaluer leur performance. Enfin, nous sélectionnerons la méthode la plus performante et l'utiliserons pour effectuer les prévisions sur la période spécifiée.

## II. Revue méthodologique

Cette section présente une revue des méthodes de prévision des séries temporelles, explorant les principaux modèles et critères de mesure utilisés pour évaluer leur performance. Nous aborderons les techniques traditionnelles telles que les modèles ARMA, ARIMA, et le lissage exponentiel, ainsi que des approches plus récentes comme Prophet et les réseaux de neurones, pour offrir une compréhension des outils disponibles pour la prévision.

### 2.1. ARMA, ARIMA et SARIMA

#### ARMA (Auto Regressive Moving Average)

Le modèle ARMA est une combinaison des approches autorégressives (AR) et de moyenne mobile (MA). Ce modèle est particulièrement utile pour modéliser des séries temporelles qui sont stationnaires, c'est-à-dire celles sans tendance claire ou saisonnalité dominante. Le modèle ARMA peut s'écrire sous la forme :

$$ARMA(p, q) : y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Avec  $y_i$  la série temporelle à la date  $i$ ,  $\phi_i$  les coefficients autorégressifs,  $\theta_i$  les coefficients moyenne mobile et  $\varepsilon_i$  l'innovation à la date  $i$ .

#### ARIMA (Auto Regressive Integrated Moving Average)

Le modèle ARIMA étend l'ARMA pour inclure la non-stationnarité en introduisant un terme de différenciation,  $d$ , qui aide à rendre la série temporelle stationnaire avant d'appliquer les effets AR et MA.

#### SARIMA (Seasonal ARIMA)

SARIMA, ou ARIMA saisonnier, noté **ARIMA (p, d, q) (P, D, Q)<sub>s</sub>** est conçu pour capturer à la fois les non-stationnarités et les saisonnalités dans les données de séries temporelles. En plus des paramètres non saisonniers  $p$ ,  $d$ , et  $q$ , SARIMA ajoute trois termes saisonniers  $P$ ,  $D$  et  $Q$  qui correspondent respectivement à l'autoregression saisonnière, à la différenciation saisonnière, et à la moyenne mobile saisonnière. Le paramètre ' $s$ ' représente la longueur de la saison (par exemple, 12 pour des données mensuelles avec une saisonnalité annuelle). Ce

modèle est extrêmement utile pour les séries qui présentent des motifs clairs et répétitifs sur une période annuelle ou une autre période spécifique.

La sélection du modèle le plus approprié pour la prévision des séries temporelles dépend fortement des caractéristiques spécifiques des données analysées. Un modèle ARMA peut être idéal pour des séries qui sont stationnaires, tandis que les modèles ARIMA et SARIMA sont préférables pour traiter des données présentant respectivement des tendances et des saisonnalités. Une fois le modèle de base choisi, l'utilisation des critères d'information tels que le critère d'information d'Akaike (AIC) ou le critère d'information bayésien (BIC) permet de déterminer les paramètres optimaux ( $p, d, q$ ) et ( $P, D, Q, S$ ) en minimisant la complexité du modèle tout en maximisant la qualité de l'ajustement.

## **2.2. Lissage exponentiel : Simple, Double, Triple**

Le lissage exponentiel est une approche populaire pour la prévision des séries temporelles, caractérisée par sa capacité à modéliser les données avec ou sans tendance et saisonnalité. Ces méthodes attribuent un poids décroissant aux observations passées, donnant plus d'importance aux données les plus récentes.

### **Lissage Exponentiel Simple (SES)**

Le lissage exponentiel simple est utilisé pour les séries sans tendance ni saisonnalité. Ce modèle ajuste le niveau de la série à chaque période, rendant la prévision équivalente à la dernière observation ajustée.

### **Lissage Exponentiel Double (DES)**

Le lissage exponentiel double, également connu sous le nom de modèle de Holt, introduit un composant de tendance en plus du niveau. Il permet de capturer les tendances linéaires dans les données, offrant ainsi une prévision plus précise pour les séries avec tendance mais sans saisonnalité. Les paramètres  $\alpha$  et  $\beta$  contrôlent respectivement la vitesse d'ajustement du niveau et de la tendance.

### **Lissage Exponentiel Triple (TES)**

Le lissage exponentiel triple, ou méthode de Holt-Winters, ajoute un composant saisonnier aux composants de niveau et de tendance. Ce modèle est idéal pour les

séries avec des motifs saisonniers prononcés. Il peut gérer différents types de saisonnalité (additive ou multiplicative), selon la nature des variations saisonnières observées dans les données. La longueur de la saison, souvent notée par 'L', est définie selon le cycle saisonnier observé, par exemple annuel ( $L=12$  pour des données mensuelles).

### **2.3. Prophet**

Prophet est une méthode de prévision robuste et flexible développée par Facebook, idéale pour les données avec des tendances fortes et des saisonnalités multiples. Ce modèle peut traiter efficacement des irrégularités dues à des vacances et des événements spéciaux, ainsi que les changements de tendance et les données manquantes. Il permet de modéliser les séries temporelles en utilisant soit un modèle additif, soit un modèle multiplicatif, ce qui le rend adapté à une grande variété de séries temporelles. Le modèle intègre trois composantes principales pour la prévision : la tendance ( $g(t)$ ), la saisonnalité ( $s(t)$ ), et les effets des vacances ou événements spéciaux ( $h(t)$ ). Ces composantes sont ajustées pour s'adapter précisément aux particularités des données.

### **2.4. Réseaux de Neurones Artificiels**

Les réseaux de neurones artificiels sont des outils de prévision puissants qui imitent la façon dont le cerveau humain fonctionne. Ils sont utilisés pour comprendre et prédire des séquences de données, comme les ventes au fil du temps. Le modèle le plus simple, appelé réseau feed-forward, comprend trois parties : une couche pour recevoir les données, des couches intermédiaires qui traitent ces informations, et une couche finale qui fait la prédiction. Pour des prévisions plus complexes impliquant des données du passé, on utilise des réseaux neuronaux récurrents (RNN). Cependant, les RNN peuvent avoir du mal à se souvenir d'informations très anciennes. Les réseaux LSTM sont une amélioration des RNN et sont conçus pour mieux se souvenir de ces informations sur de longues périodes, ce qui les rend très utiles pour prédire des événements futurs basés sur des tendances historiques longues. Parmi les méthodes disponibles, nous avons les réseaux feed-forward, les réseaux neuronaux récurrents (RNN), et les réseaux à mémoire longue (LSTM).

Pour présenter les mesures d'erreur de prévision utilisées pour évaluer les modèles de séries temporelles, voici un résumé basé sur les informations fournies dans l'image :

### Mesures d'Erreur de Prévision pour les Séries Temporelles

L'évaluation des performances des modèles de séries temporelles est cruciale pour vérifier leur précision et leur applicabilité. Trois mesures d'erreur courantes sont utilisées pour cette évaluation :

➤ Mean Squared Error (MSE)

Il s'agit de la moyenne des carrés des différences entre les valeurs observées et les valeurs prédites. Cette mesure est sensible aux valeurs aberrantes car elle élève les erreurs au carré.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Avec  $Y_i$  la valeur réelle et  $\hat{Y}_i$  la valeur prédite pour l'observation  $i$  et  $n$  le nombre total d'observations.

➤ Root Mean Squared Error (RMSE)

Il s'agit simplement de la racine carrée du MSE, ce qui rend cette mesure dans les mêmes unités que les données d'origine, facilitant ainsi leur interprétation.

$$RMSE = \sqrt{MSE}$$

➤ Mean Absolute Percentage Error (MAPE)

Le MAPE exprime l'erreur en pourcentage, ce qui permet de comparer facilement la performance entre différents jeux de données ou modèles. Il calcule la moyenne des valeurs absolues des erreurs en pourcentage par rapport aux valeurs réelles.

$$MAPE = \left( \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \right) \times 100$$

Au terme de cette revue méthodologique, nous avons exploré une variété de modèles de prévision des séries temporelles, ainsi que des critères de mesure de performance. Nous ne prévoyons pas d'utiliser les techniques de Machine Learning telles que les réseaux de neurones. La sélection des modèles appropriés sera déterminée après une analyse approfondie des caractéristiques spécifiques de nos données dans les sections ultérieures. Cette étape nous permettra de choisir les méthodes les plus adaptées, que nous évaluerons ensuite à l'aide des critères de mesure établis. Ce processus garantira que notre prévision finale soit réalisée avec la plus grande précision et efficacité possible.



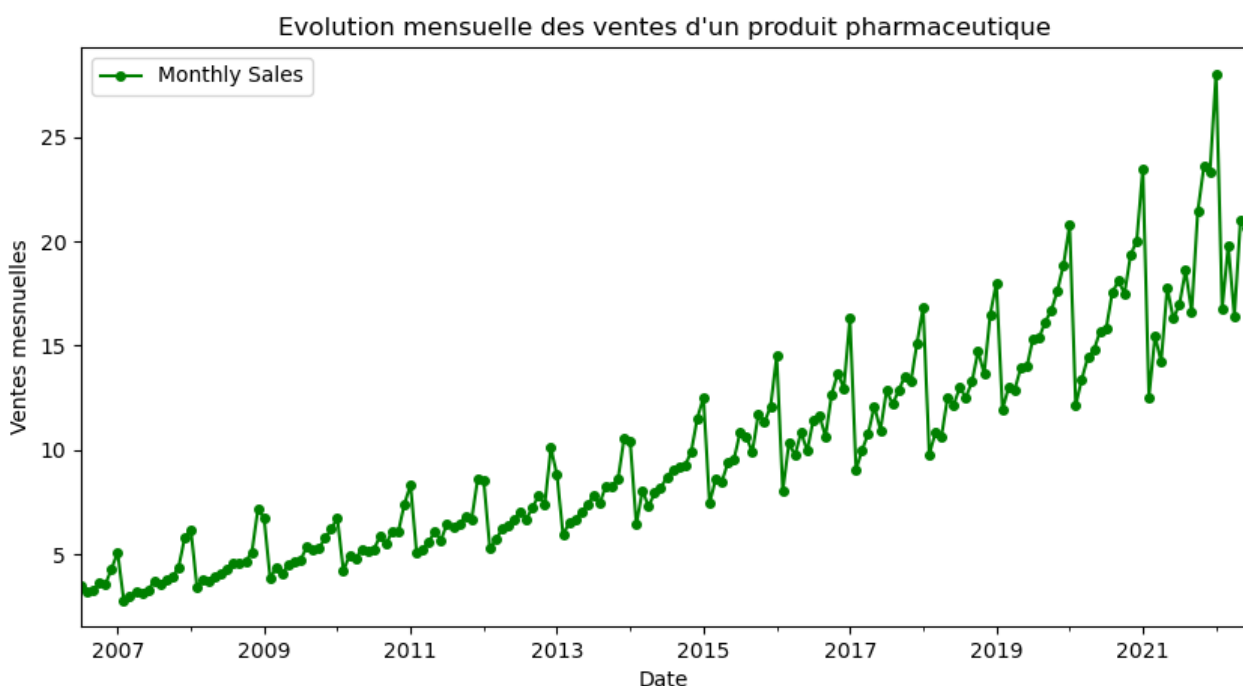
## III. Exploration des données

### 3.1. Valeurs manquantes

Nous avons analysé les ventes mensuelles de notre série temporelle. La première étape de notre démarche a consisté à vérifier la qualité des données. Nous avons ainsi constaté qu'il n'y avait aucune valeur manquante dans notre ensemble de données. Cette absence de valeurs manquantes nous a permis de procéder directement à l'analyse sans avoir à traiter des données manquantes, ce qui est souvent une source de biais ou d'incertitude dans les prévisions.

### 3.2. Visualisation de la série

Le graphique ci-dessous illustre l'évolution des ventes mensuelles au fil du temps. On observe une tendance générale à la hausse des ventes, indiquant une augmentation progressive de la demande pour le produit pharmaceutique étudié. En outre, les ventes montrent un mouvement cyclique croissant à partir de février, avec un pic très marqué au début de chaque année, suivi d'une légère décroissance avant de reprendre une tendance haussière. Cette fluctuation saisonnière suggère une saisonnalité dans les données, un aspect crucial à prendre en compte pour des prévisions précises.



Par ailleurs, le jeu de données semble révéler un schéma multiplicatif. En effet, l'amplitude des variations semble croître au fur et à mesure du temps, ce qui suggère que le schéma additif ne capterait pas correctement la structure de saisonnalité, laissant les résidus autocorrélés. Les caractéristiques de cette tendance et de cette saisonnalité seront examinées plus en détail dans la section suivante.

### **3.3. Analyse de la Tendance et de la Saisonnalité**

L'analyse des composantes de la série temporelle, à savoir la tendance et la saisonnalité, est importante pour plusieurs raisons. Premièrement, elle permet de comprendre les comportements sous-jacents de la série, comme les augmentations ou diminutions à long terme (tendance) et les fluctuations régulières (saisonnalité). Cette compréhension est essentielle pour des prévisions précises et pour prendre des décisions éclairées basées sur les données. Deuxièmement, la détection et la modélisation correctes de ces composantes améliorent la précision des modèles de prévision en réduisant les résidus et en capturant mieux les patterns de la série.

#### **Exploration visuelle des composantes de la série temporelle**

Passons maintenant à explorer les méthodes de décomposition, qui sont des outils essentiels pour analyser les tendances et la saisonnalité des séries temporelles.

Les méthodes de décomposition de série temporelle permettent de séparer une série en ses différentes composantes : tendance, saisonnalité et résidu. Les principales méthodes incluent :

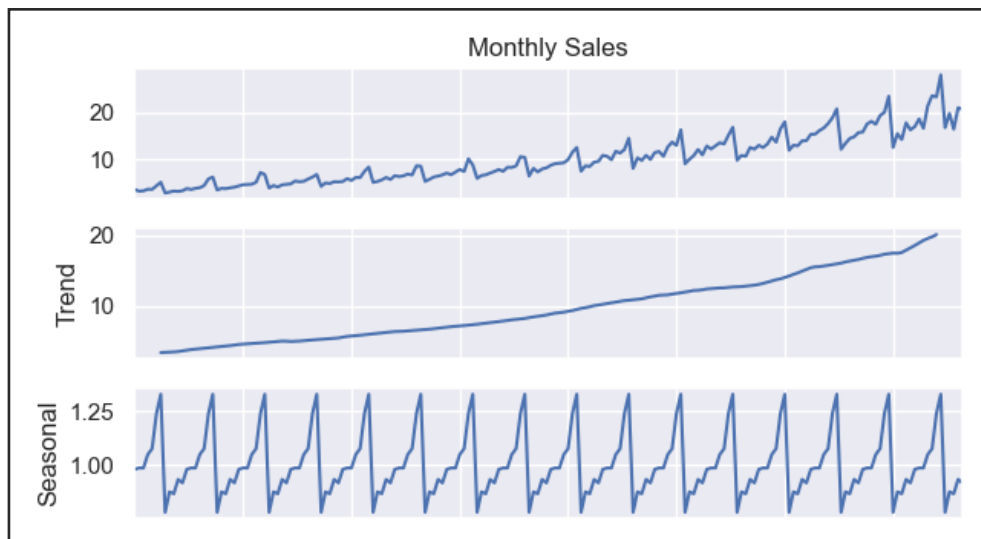
##### **➤ Décomposition Additive**

La série est décomposée en la somme des composantes (tendance + saisonnalité + résidu). Elle est utilisée lorsque la saisonnalité a une amplitude constante au fil du temps.

##### **➤ Décomposition Multiplicative**

La série est décomposée en le produit des composantes (tendance \* saisonnalité \* résidu). Elle est utilisée lorsque l'amplitude de la saisonnalité change proportionnellement à la tendance.

Pour une meilleure appréhension des dynamiques sous-jacentes de notre série de données, nous allons approfondir l'interprétation des résultats de la décomposition multiplicative. À cet effet, nous avons représenté graphiquement les différentes composantes en choisissant la décomposition multiplicative étant donné qu'avec la visualisation graphique on observait une augmentation de l'amplitude de la série.

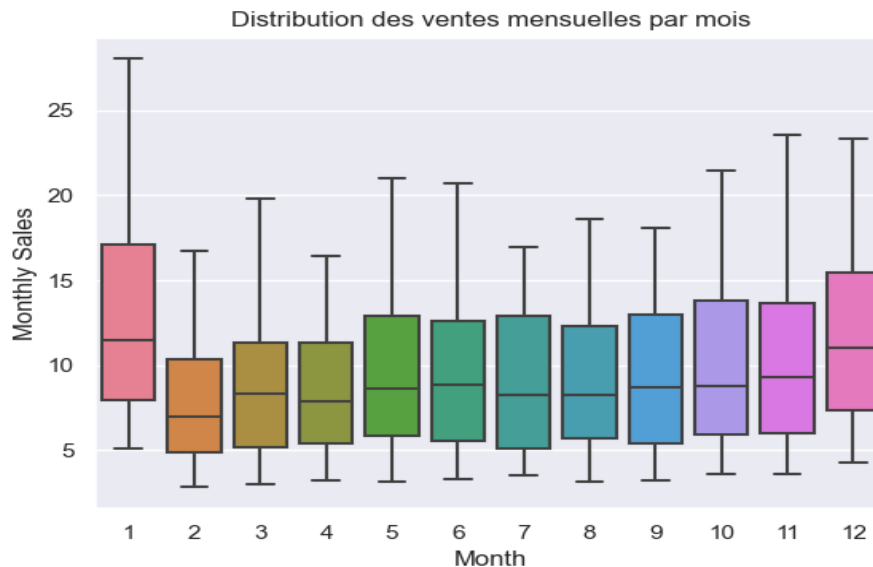


Le graphique ci-dessus montre clairement une tendance croissante et une saisonnalité marquée.

- **Trend** (Tendance) : Cette composante montre une augmentation progressive des ventes au fil du temps. Cela confirme que la demande pour le produit pharmaceutique étudié augmente régulièrement.
- **Seasonal** (Saisonnalité) : La composante saisonnière révèle des pics réguliers en début d'année. Cela indique une forte saisonnalité dans les ventes, où certaines périodes de l'année voient des ventes beaucoup plus élevées que d'autres.

### Identification de la saisonnalité

Nous avons enrichi notre analyse de la saisonnalité en complétant notre analyse avec un boxplot.



Les Boxplots mettent en évidence une tendance claire : les ventes mensuelles du produit pharmaceutique atteignent leur apogée en janvier. De manière générale les niveaux de la distribution varient selon les mois ce qui nous indique la présence de saisonnalité mensuelles.

#### Test de Kendall pour la tendance

Le test de Kendall, ou test de tendance de Kendall, est une méthode non paramétrique utilisée pour déterminer la présence d'une tendance monotone dans une série temporelle. Ce test est particulièrement utile pour vérifier si les observations augmentent ou diminuent systématiquement au fil du temps sans faire d'hypothèses sur la distribution des données.

Les hypothèses de ce test sont les suivantes :

$$\begin{cases} H_0 : \text{Il n'y a pas de tendance dans les données} \\ H_1 : \text{Il existe une tendance monotone dans les données (soit croissante, soit décroissante)} \end{cases}$$

Ce test calcule un coefficient appelé tau de Kendall ( $\tau$ ), qui mesure la force et la direction de l'association entre deux variables. Les valeurs de  $\tau$  varient de -1 à 1, où :

- Une valeur de 1 indique une tendance monotone strictement croissante.
- Une valeur de -1 indique une tendance monotone strictement décroissante.
- Une valeur de 0 indique l'absence de tendance monotone.

L'application de ce test à notre série temporelle nous permettra de confirmer statistiquement la tendance croissante observée visuellement et de renforcer la robustesse de nos prévisions.

Passons maintenant à l'application pratique du test de Kendall sur nos données de ventes mensuelles pour vérifier la tendance observée.

Test du Tau de Kendall :

<b>Tau de Kendall</b>	<b>P-Value</b>
0.837	1,33E-66

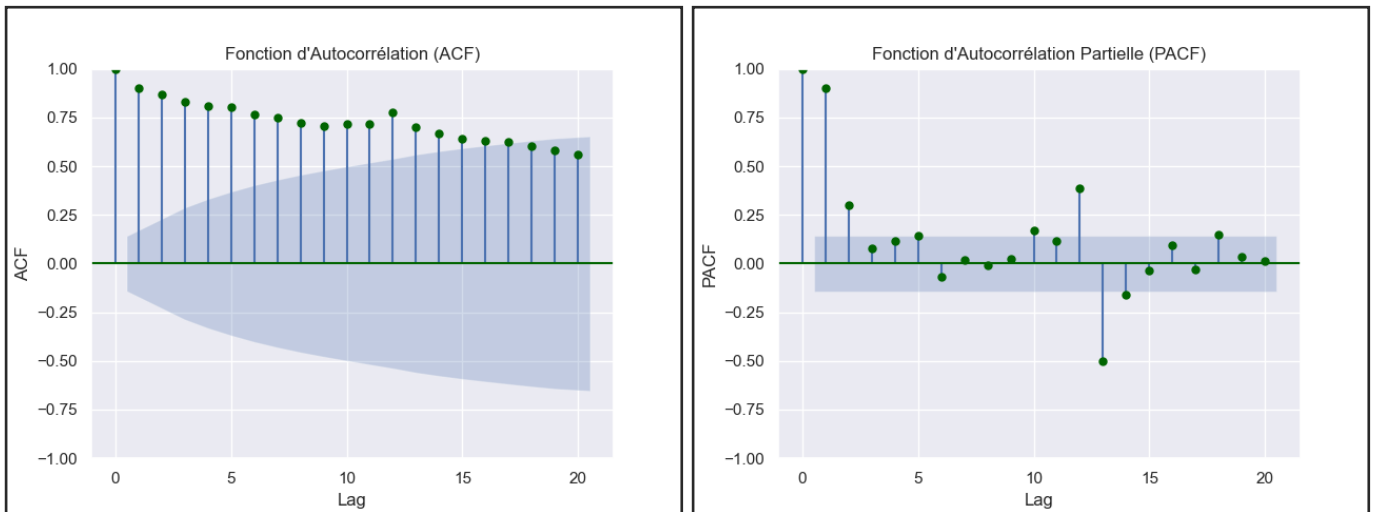
En conséquence, nous pouvons confirmer qu'il existe une tendance monotone significative dans les ventes mensuelles, ce qui corrobore nos observations visuelles précédentes.

### **3.4. Analyse de la stationnarité**

La stationnarité est un concept important en analyse de séries temporelles. Une série temporelle est dite stationnaire si ses propriétés statistiques, telles que la moyenne, la variance et l'autocorrélation, restent constantes dans le temps. En d'autres termes, les caractéristiques de la série ne changent pas, ce qui permet de prévoir plus facilement les valeurs futures en se basant sur les données historiques.

L'analyse de la stationnarité est importante car la plupart des méthodes de prévision, notamment les modèles ARIMA (AutoRegressive Integrated Moving Average), supposent que les séries temporelles utilisées sont stationnaires. Si la série n'est pas stationnaire, il est souvent nécessaire de la transformer pour la rendre stationnaire avant d'appliquer ces modèles.

Pour débiter nos analyses nous allons d'abord examiner les fonctions d'autocorrélation (ACF) et fonctions d'autocorrélation partielle (PACF) pour nous faire une idée sur les caractéristiques de la série.



L'analyse de la fonction d'autocorrélation (ACF) montre une décroissance très lente de l'autocorrélation, typique des séries temporelles non stationnaires. Par ailleurs, la fonction d'autocorrélation partielle (PACF) présente un pic notable au décalage de 12 mois, signalant une corrélation significative à ce retard spécifique. Ensemble, ces éléments indiquent que la série des ventes est non stationnaire et qu'elle exhibe une saisonnalité annuelle.

Après avoir exploré visuellement les composantes de la série temporelle, nous devons confirmer ces observations par des méthodes statistiques rigoureuses.

### Test de Dickey-Fuller

Pour vérifier la stationnarité de notre série temporelle, nous utilisons le test de Dickey-Fuller (ADF), qui est un test statistique utilisé pour vérifier la présence d'une racine unitaire dans une série temporelle. La présence d'une racine unitaire indique que la série est non stationnaire.

Les hypothèses du test de Dickey-Fuller sont les suivantes :

- **Hypothèse nulle (H0)** : La série temporelle possède une racine unitaire (elle est non stationnaire).
- **Hypothèse alternative (H1)** : La série temporelle ne possède pas de racine unitaire (elle est stationnaire).

En appliquant le test de Dickey-Fuller à nos données de ventes mensuelles, nous avons trouvé une statistique de test égale à 3.06 et une p-value égale à 1. Ces résultats indiquent que la série n'est pas stationnaire, ce qui était prévisible.

Après avoir conclu quant à la présence d'une tendance et d'une saisonnalité dans notre série temporelle, cela oriente nos choix concernant le type de modèle à utiliser. En effet, en présence de saisonnalité, des modèles tels que SARIMA, Prophet ou le TES (Triple Exponential Smoothing) peuvent être appropriés pour modéliser la série temporelle.

## IV. Application des modèles et résultats

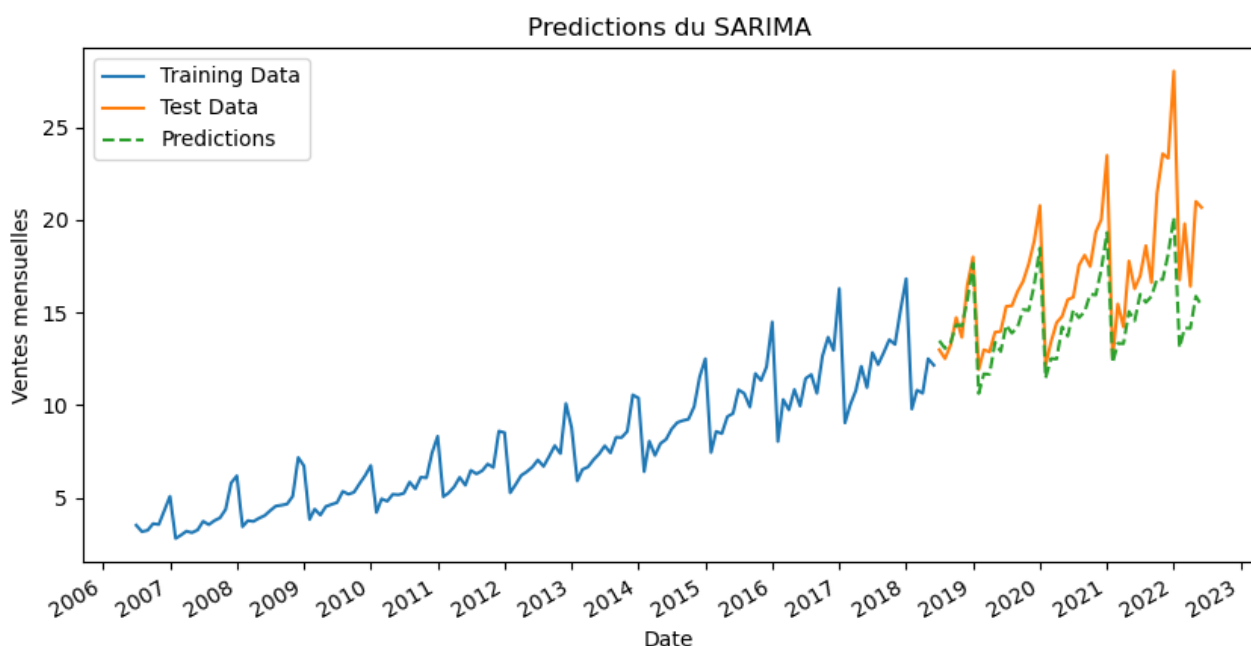
Cette partie est dédiée à présenter les résultats des modèles estimés afin de déterminer quel est le meilleur candidat pour prédire les ventes. Pour ce faire, nous avons scindé les données disponibles en 2 parties :

L'échantillon d'apprentissage (**07-2006 au 06-2018**) qui va permettre d'entraîner les modèles et l'échantillon de test (**07-2018 au 06-2022**). Les données de test vont servir à confronter l'erreur de prédiction des différents modèles afin de déterminer lequel affine le mieux la structure de ventes du produit pharmaceutique.

### 4.1. SARIMA

Le modèle SARIMA à différence des modèles ARMA et ARIMA est qu'il permet d'intégrer une composante de saisonnalité au sein du modèle. C'est pour cette raison que nous présentons uniquement le SARIMA et non pas les autres. Nous avons utilisé la fonction `auto_arima` du package `pmdarima` qui permet de choisir l'ordre optimal au sens d'un critère d'information. Nous avons choisi le critère AIC et celui-ci a révélé comme modèle :  $SARIMA(0,1,2)(0,1,1)[12]$ . Remarquons uniquement que nous avons imposé une saisonnalité de 12 périodes ce qui correspond à une année puisque nous disposons de données mensuelles.

Voici les résultats du modèle :



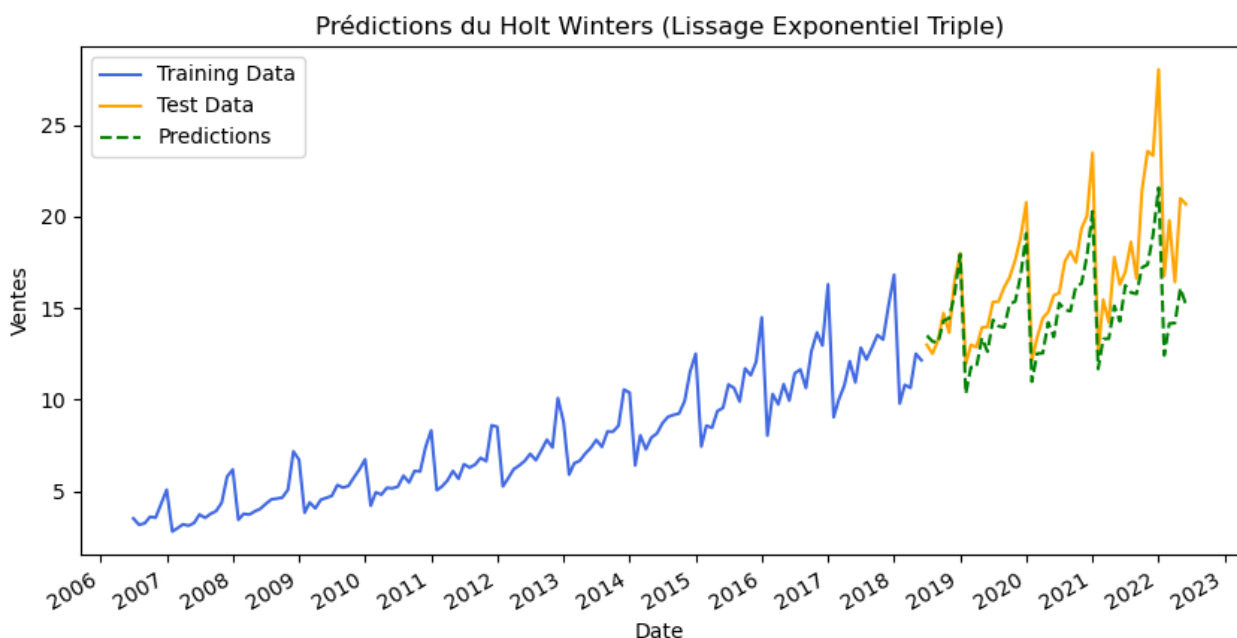


Visuellement, on remarque que le modèle SARIMA semble approcher la bonne structure de saisonnalité mais n'arrive pas à approcher correctement la tendance de la série.

## 4.2. TES (Triple Exponential Smoothing)

La méthode de lissage exponentiel triple, aussi connue sous le nom de méthode de Holt-Winters, est particulièrement efficace pour les données ayant des motifs saisonniers clairs et répétitifs. Nous avons paramétré le TES en lui imposant un schéma de saisonnalité multiplicatif comme nous avons pu identifier en premier lieu. De plus, il comprend une composante de niveau et de tendance.

Voici les résultats du modèle :



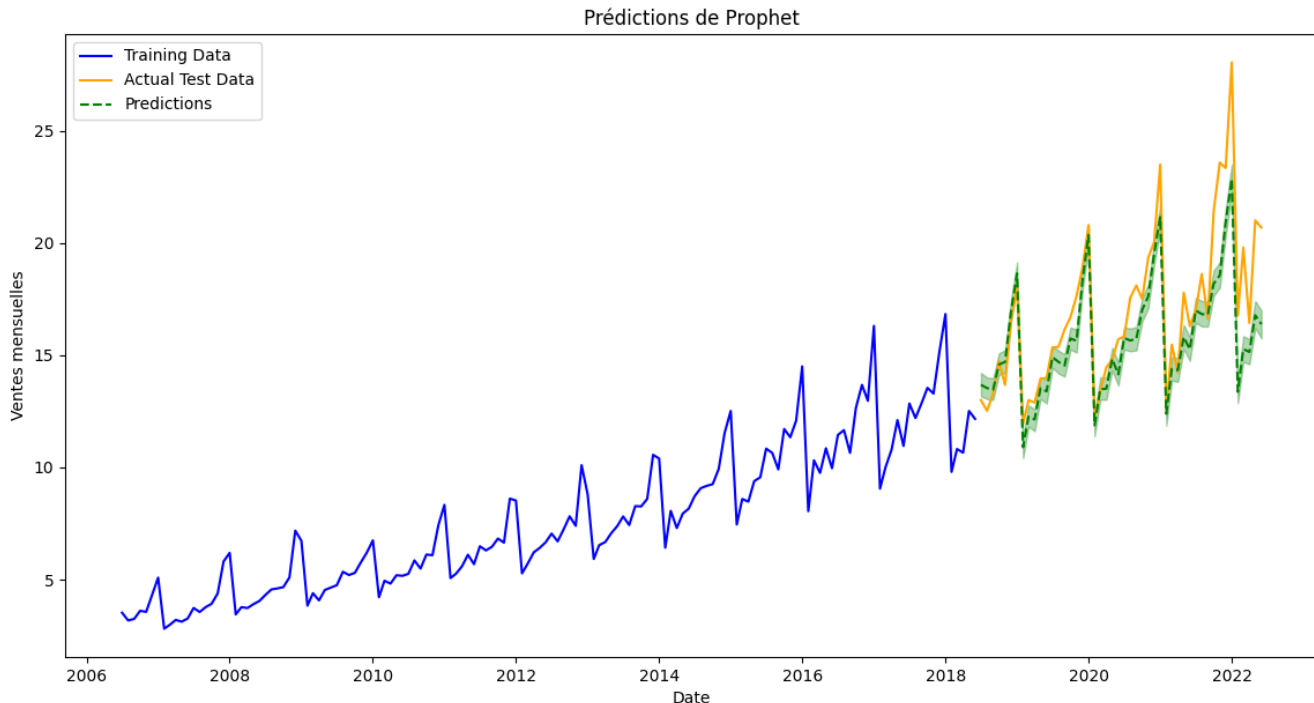
On remarque de manière plutôt claire que le modèle de Holt Winters permet de mieux approcher la série, en particulier pendant la dernière année. En effet, la différence de prévision et des données observées pour le dernier pic en 2022 se rapprochent de plus en plus.

## 4.3. Prophet

Le modèle Prophet qui a été développé en Open Source par Facebook permet également d'approcher la composante saisonnière et la composante de tendance. Bien qu'il puisse être paramétré pour indiquer des jours fériés ou encore être

robuste face à des valeurs manquantes, nous avons gardé sa forme la plus simple en précisant un schéma de saisonnalité multiplicatif.

Voici les résultats de ce modèle :



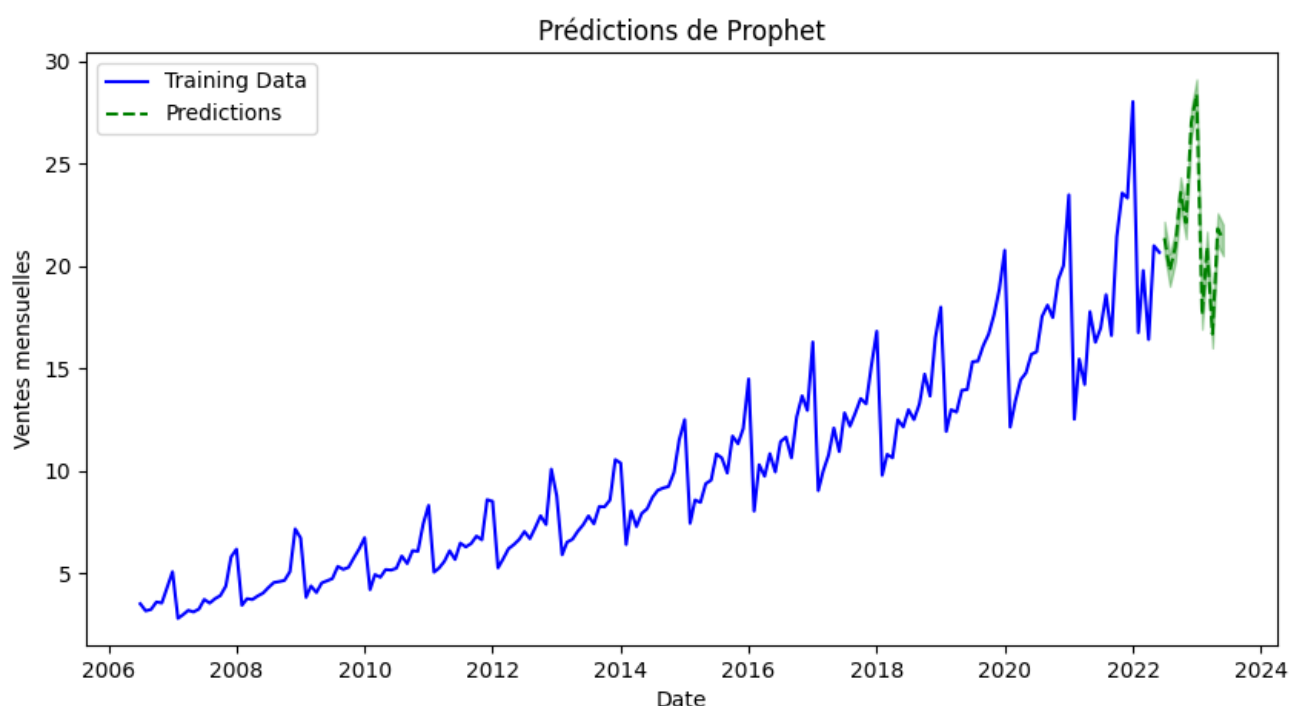
Nous observons que les prévisions fournies par le modèle Prophet semblent significativement meilleures que celles fournis par le TES.

#### 4.4. Prévion finale

Afin de confronter ces 3 modèles, nous présentons ci-dessous un tableau présentant les métriques citées précédemment (MSE RMSE MAPE) afin de déterminer celle qui approche le mieux la série de données décrivant la vente du produit pharmaceutique.

Modèles	MSE	RMSE	MAPE
SARIMA	8.070837	2.840922	0.142392
TES	7.058989	2.656876	0.115305
Prophet	3.835433	1.958426	0.074688

On remarque de manière plutôt évidente qu'il s'agit du modèle Prophet qui minimise les 3 critères décrivant l'erreur de prévision, indiquant qu'il s'agit du modèle le plus performant. Notons que nous avons essayé d'appliquer la fonction *OnlineEnsembleForecaster* qui permet de choisir des poids optimaux pour plusieurs prédicteurs au prorata de la MSE. Toutefois, celui-ci ne présentait pas des meilleures métriques que celles présentés par le modèle Prophet. Pour conclure, nous décidons d'utiliser le modèle Prophet afin de prédire nos données puisque c'est celui qui présente les meilleures performances.



Les valeurs détaillées de ces prévisions finales pour la période 07-2022 à 06-2023 sont fournies en annexe.

## V. Conclusion

Notre projet a évalué plusieurs modèles de prévision des séries temporelles pour les ventes d'un produit pharmaceutique, en utilisant les méthodes SARIMA, TES et Prophet. Nos données, qui couvrent la période de 2006 à 2022, présentaient une non-stationnarité avec des tendances et une saisonnalité marquée. Après une analyse comparative approfondie sur un échantillon de test, nous avons constaté que le modèle Prophet surpasse les autres avec les meilleures valeurs de MSE, RMSE, et MAPE.

Malgré les tentatives d'améliorer ces prédictions en utilisant des techniques combinées de plusieurs modèles, Prophet a été retenu pour sa simplicité et sa performance supérieure. Les prévisions obtenues avec Prophet montrent des variations saisonnières claires et correspondent bien aux tendances attendues du marché, suggérant des pics de ventes pendant les mois d'hiver.

En conséquence, nous avons choisi Prophet pour réaliser la prédiction finale des ventes du produit pharmaceutique de juillet 2022 à juin 2023. Ces prévisions permettent de recommander des ajustements stratégiques pour la production et la gestion des stocks, afin de répondre efficacement à la demande future.

## Annexe

### Tableau des prévisions

Date	Prévisions des ventes
2022/07/01	21.376638
2022/08/01	19.780929
2022/09/01	20.941299
2022/10/01	23.599939
2022/11/01	22.132601
2022/12/01	27.005992
2023/01/01	28.394001
2023/02/01	17.639181
2023/03/01	20.944374
2023/04/01	16.694010
2023/05/01	21.831088
2023/06/01	21.307947