

PRÉVOIR LES VENTES AVEC DES SÉRIES TEMPORELLES

Utilisation des séries temporelles pour prédire les ventes
et optimiser les stocks



PRÉSENTATION



Dans un marché en constante évolution, l'anticipation des ventes est cruciale pour les entreprises. Une mauvaise gestion des stocks peut provoquer des ruptures ou des surplus coûteux.

L'objectif de ce projet est de développer un modèle capable de prévoir les ventes à l'aide de techniques de séries temporelles, afin d'optimiser les ressources et d'améliorer la prise de décision.

Ce projet s'inscrit dans un contexte réel où les données historiques des ventes d'une chaîne de magasins sont exploitées pour identifier des tendances, des saisonnalités et des variations spécifiques.

Nous exploiterons des modèles de prévision tels qu'**ARIMA** et **Prophet**, accompagnés d'outils de visualisation pour analyser et interpréter les résultats.

OBJECTIFS DU PROJET

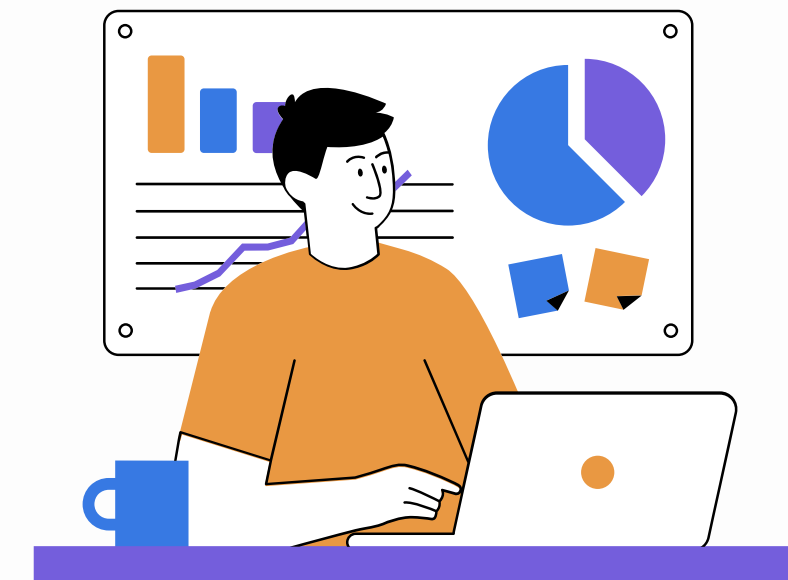
Les séries temporelles permettent de :



Identifier les
tendances et la
saisonnalité dans les
ventes.



Prédire les
performances
futures à l'aide de
modèles robustes.



Fournir des insights
exploitables pour
ajuster les stocks et
optimiser les
ressources.

BASE DE DONNÉES : STORE SALES - TIME SERIES FORECASTING

Nous utiliserons la base de données Store Sales - Time Series Forecasting qui contient plusieurs fichiers de données :

1. **train.csv**: Données historiques des ventes par magasin et catégorie de produit.
2. **test.csv**: Données sans les valeurs de ventes à prédire.
3. **transactions.csv**: Nombre de transactions par magasin et par jour.
4. **oil.csv**: Prix du pétrole, qui peut influencer les coûts et la demande.
5. **holidays_events.csv**: Jours fériés et événements spéciaux susceptibles d'affecter les ventes.
6. **stores.csv**: Informations sur chaque magasin (localisation, type, etc.).
7. **sample_submission.csv**: Format attendu pour la soumission des prédictions.



PLAN DE TRAVAIL - ÉTAPES CLÉS

I. PRÉPARATION DES DONNÉES



Chargement et exploration des données :

- Vérification des formats et types de variables.
- Gestion des valeurs manquantes et des anomalies.

Fusion des différentes sources de données :

- Intégration des informations utiles (pétrole, jours fériés, transactions).
- Transformation des dates et création de nouvelles variables pertinentes.

I. PRÉPARATION DES DONNÉES

Importation des bibliothèques

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from statsmodels.tsa.arima.model import ARIMA
from prophet import Prophet
from sklearn.metrics import mean_absolute_error,
mean_squared_error
```

I. PRÉPARATION DES DONNÉES

Chargement des données

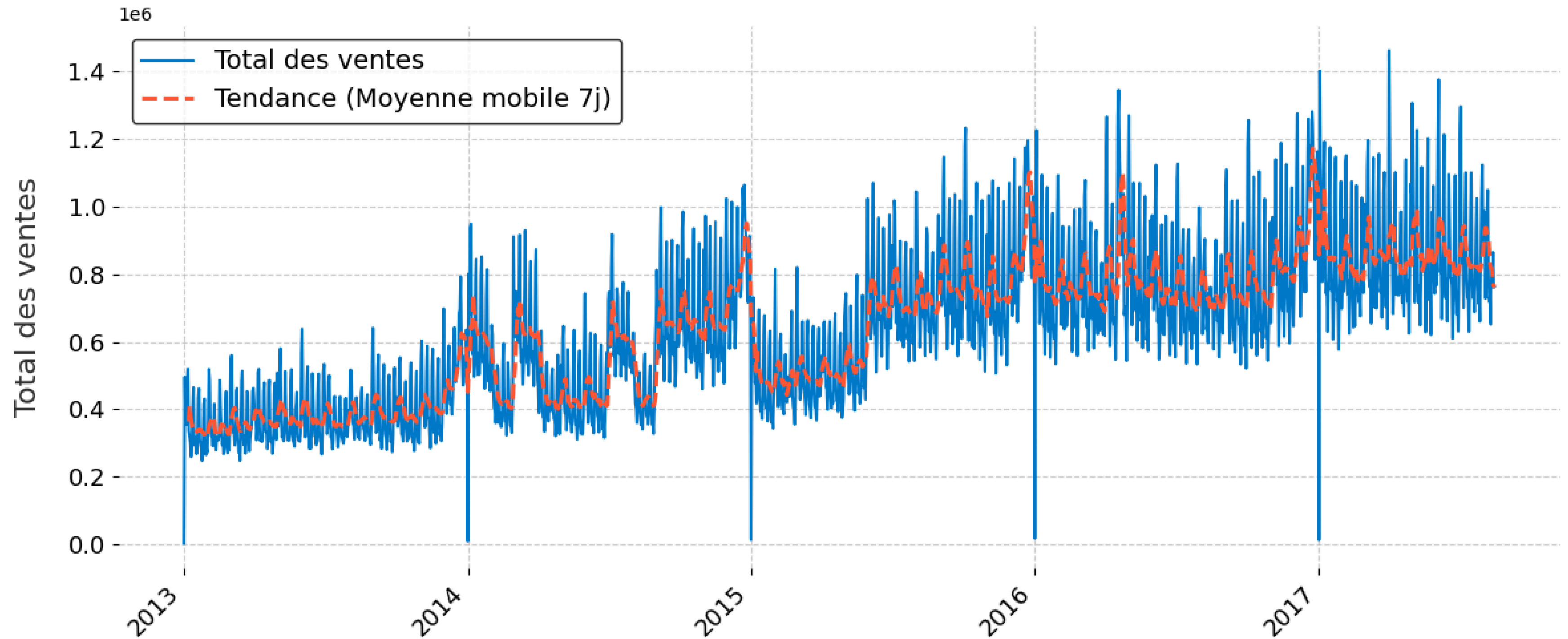
```
# Chargement des fichiers CSV
file_paths = {
    "train": "train.csv",
    "test": "test.csv",
    "transactions": "transactions.csv",
    "oil": "oil.csv",
    "holidays_events": "holidays_events.csv",
    "stores": "stores.csv",
    "sample_submission": "sample_submission.csv"
```


II. ANALYSE EXPLORATOIRE

- Visualisation des tendances globales des ventes.
- Identification de la saisonnalité et des événements impactants.
- Étude des corrélations entre les ventes et d'autres facteurs (prix du pétrole, promotions, jours fériés).

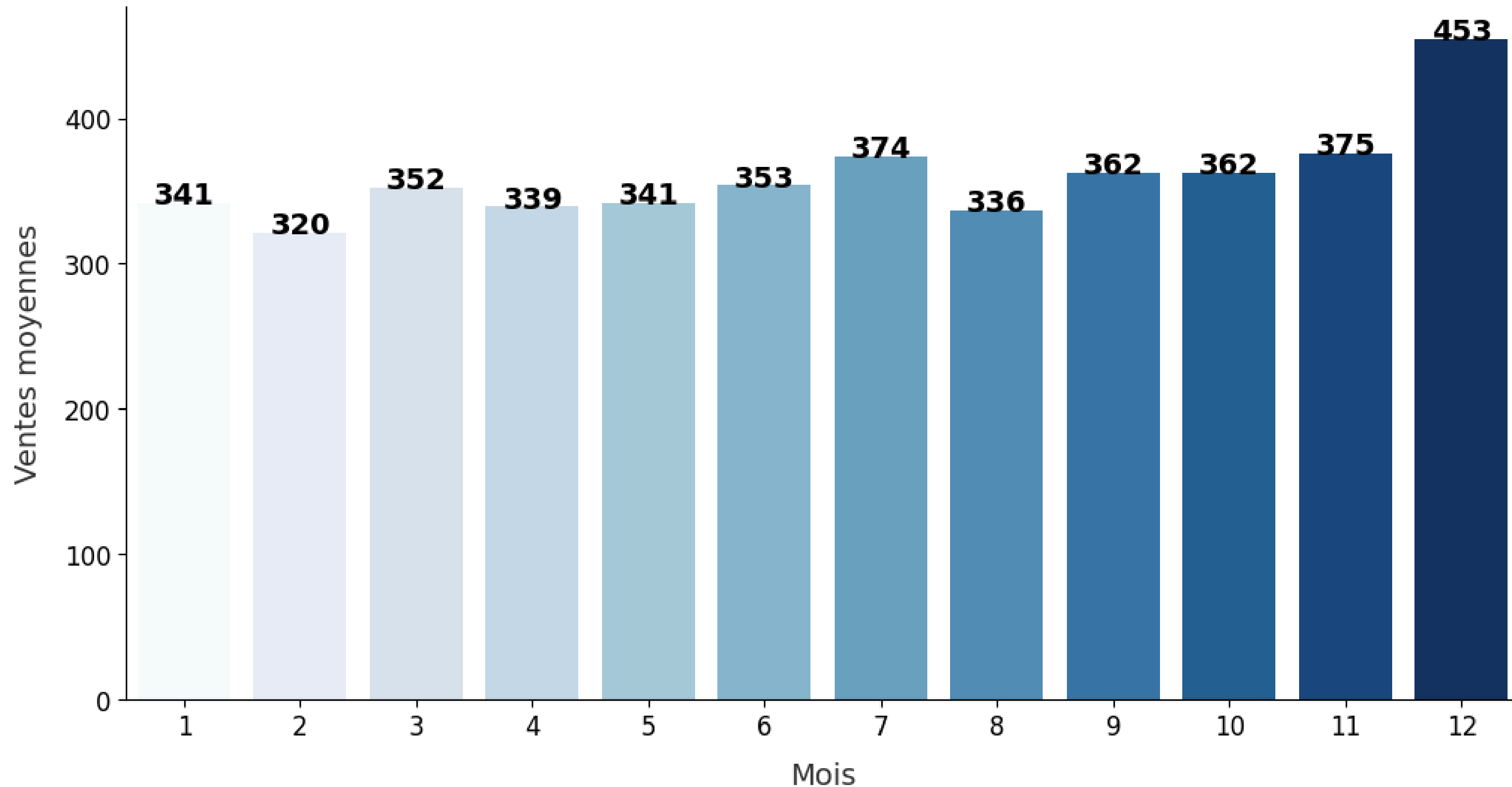
TENDANCE GLOBALE DES VENTES

ÉVOLUTION DES VENTES AU FIL DU TEMPS

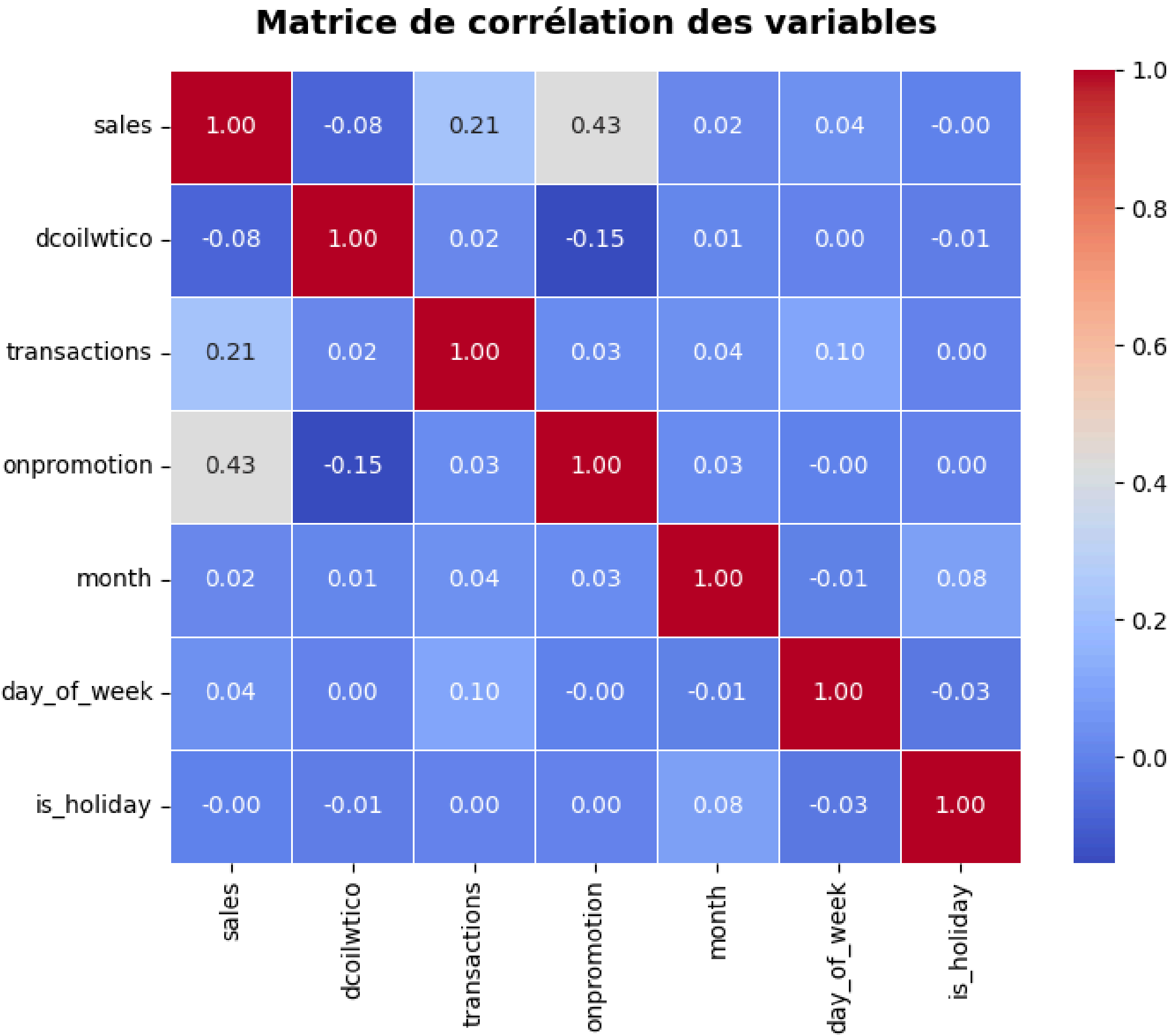


VENTES PAR MOIS POUR DÉTECTER UNE SAISONNALITÉ

Moyenne des ventes par mois



HEATMAP DES CORRÉLATIONS AVEC LES FACTEURS EXTERNES (PÉTROLE, PROMOTIONS, JOURS FÉRIÉS, TRANSACTIONS).



III. MODÉLISATION ET PRÉVISION

- **Méthodes classiques : ARIMA**

Décomposition de la série temporelle.

Sélection des paramètres optimaux (p , d , q).

Évaluation des performances.

- **Méthode avancée : Prophet**

Gestion automatique de la saisonnalité.

Prise en compte des jours fériés.

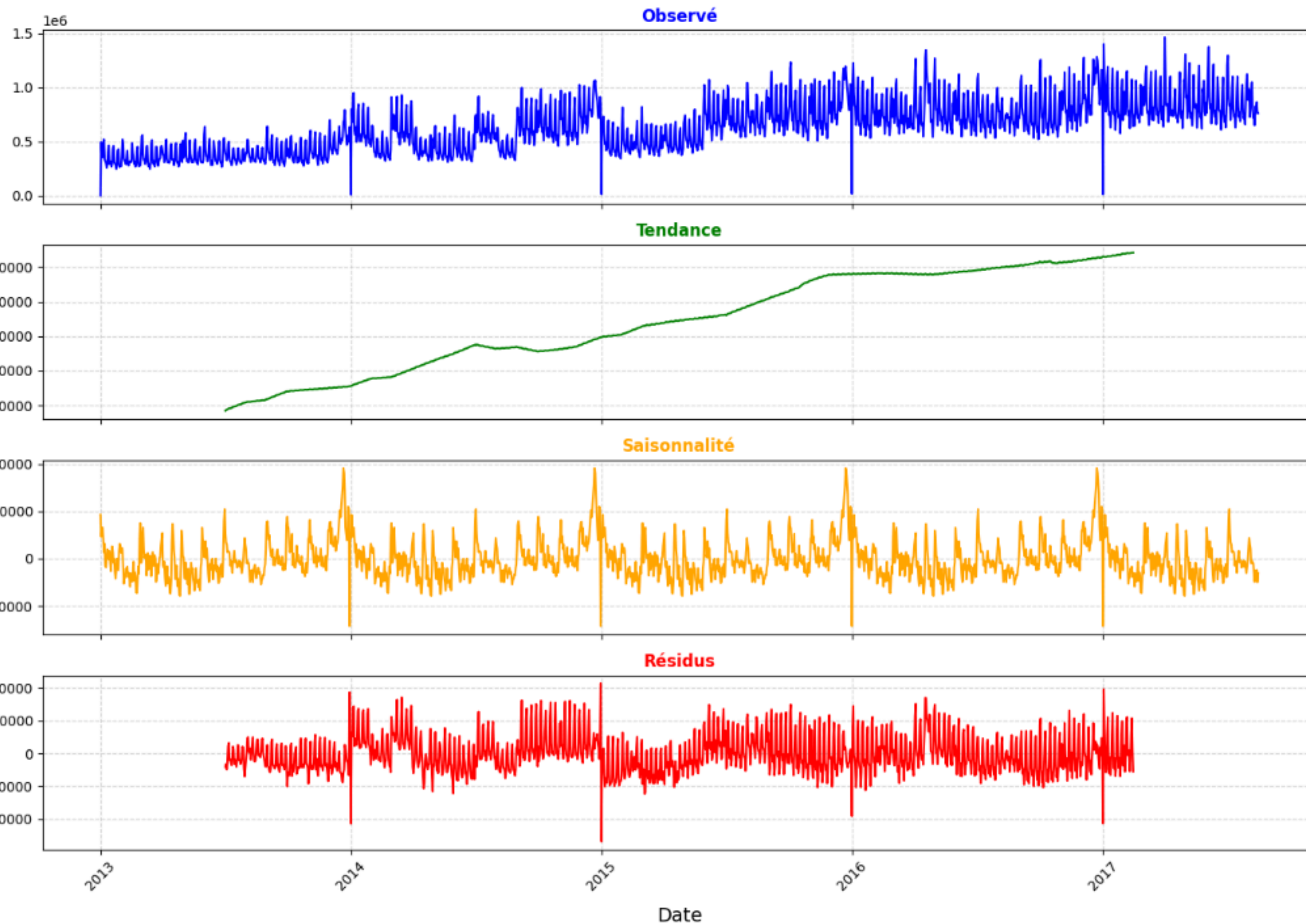
Génération des prévisions et visualisation.

Ces modèles permettent aux entreprises d'anticiper la demande, d'optimiser les stocks et d'améliorer la gestion des ressources.

MÉTHODES CLASSIQUES : ARIMA

- **L'ARIMA EST UN MODÈLE UTILISÉ POUR LA PRÉVISION DES SÉRIES TEMPORELLES.**
- **IL PERMET DE COMPRENDRE LES COMPOSANTES DE LA DEMANDE EN IDENTIFIANT LA TENDANCE, LA SAISONNALITÉ ET LES VARIATIONS ALÉATOIRES.**

DÉCOMPOSITION D'UNE SÉRIE TEMPORELLE : TENDANCE, SAISONNALITÉ ET RÉSIDUS



CETTE DÉCOMPOSITION PERMET D'ANALYSER LES TENDANCES SOUS-JACENTES DES VENTES :

◆ TENDANCE : ÉVOLUTION À LONG TERME DES VENTES.

◆ SAISONNALITÉ : EFFETS RÉPÉTITIFS LIÉS AUX PÉRIODES (EX : NOËL, ÉTÉ).

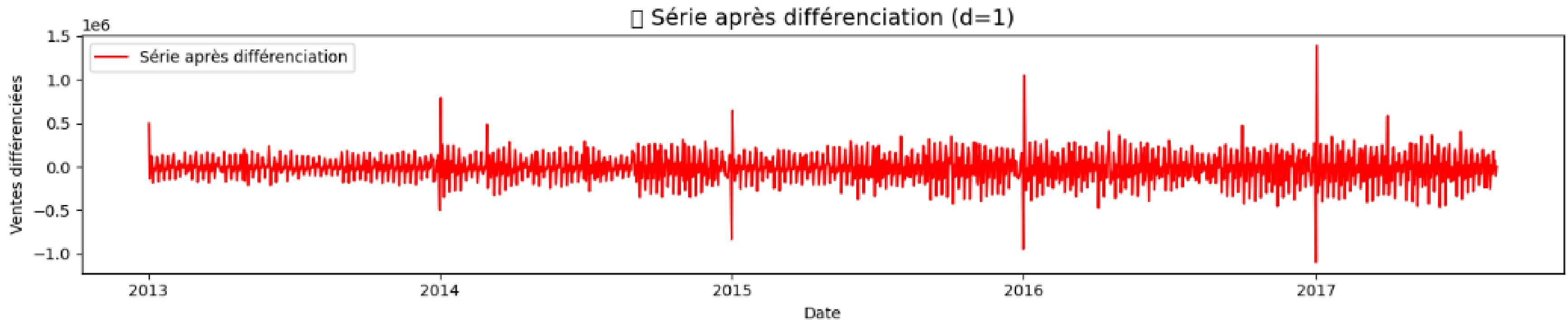
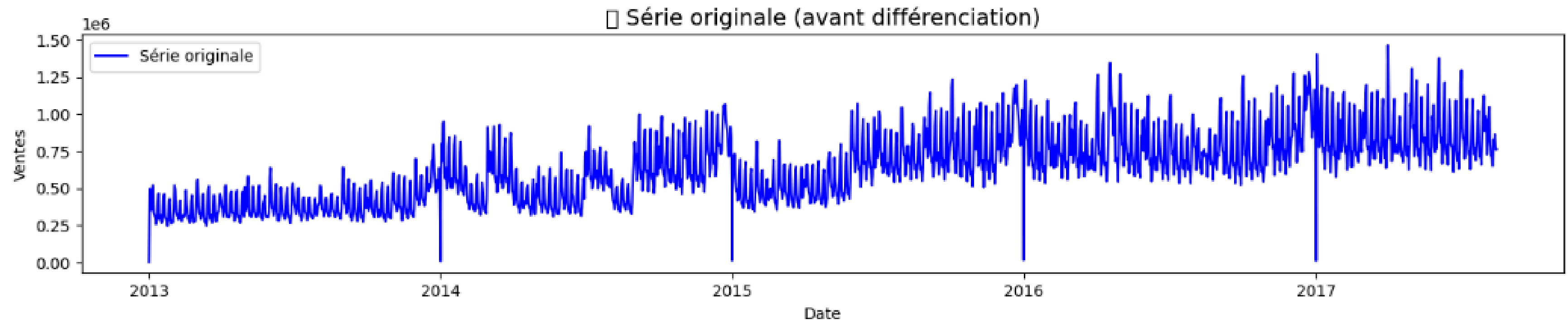
▼ RÉSIDUS : VARIATIONS ALÉATOIRES, ANOMALIES OU BRUIT DANS LES DONNÉES.

SÉLECTION DES PARAMÈTRES OPTIMAUX (P, D, Q)

1. TEST ADF (STATIONNARITÉ DE LA SÉRIE)

TEST DE STATIONNARITÉ ADF : LA SÉRIE EST-ELLE STATIONNAIRE ?

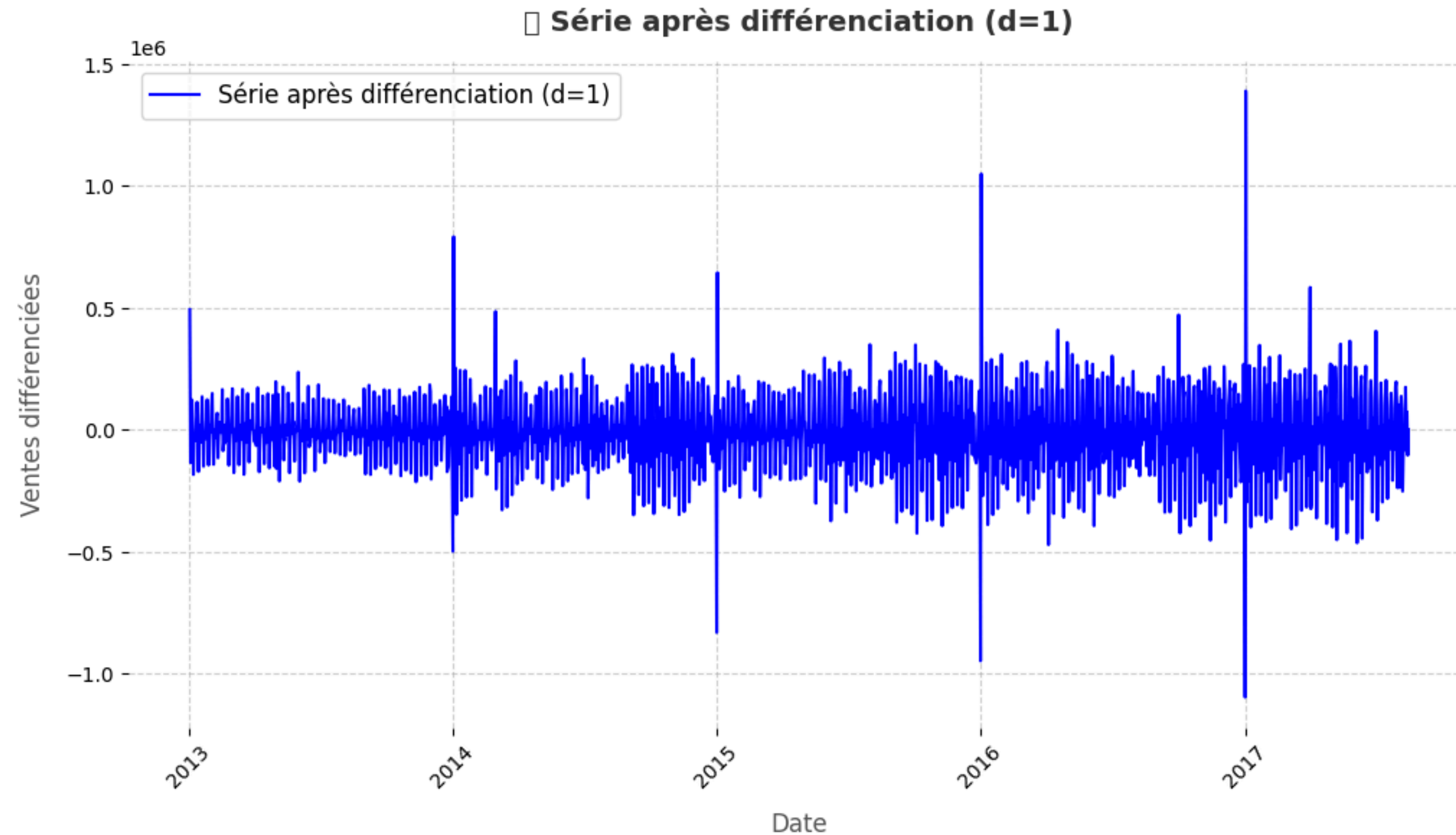
- LE TEST ADF VÉRIFIE SI LA SÉRIE TEMPORELLE EST STATIONNAIRE.
- VALEUR $P = 0,0897 > 0,05$ ✗ LA SÉRIE N'EST PAS STATIONNAIRE.
- NOUS DEVONS APPLIQUER UNE DIFFÉRENCIATION ($D=1$) POUR STABILISER LES TENDANCES.



2. TRANSFORMATION DE LA SÉRIE (DIFFÉRENCIATION D=1)

RENDRE LA SÉRIE STATIONNAIRE AVEC UNE DIFFÉRENCIATION

- UNE DIFFÉRENCIATION PERMET DE SUPPRIMER LA TENDANCE POUR STABILISER LA SÉRIE.
- APRÈS DIFFÉRENCIATION, LA SÉRIE DEVIENT STATIONNAIRE.

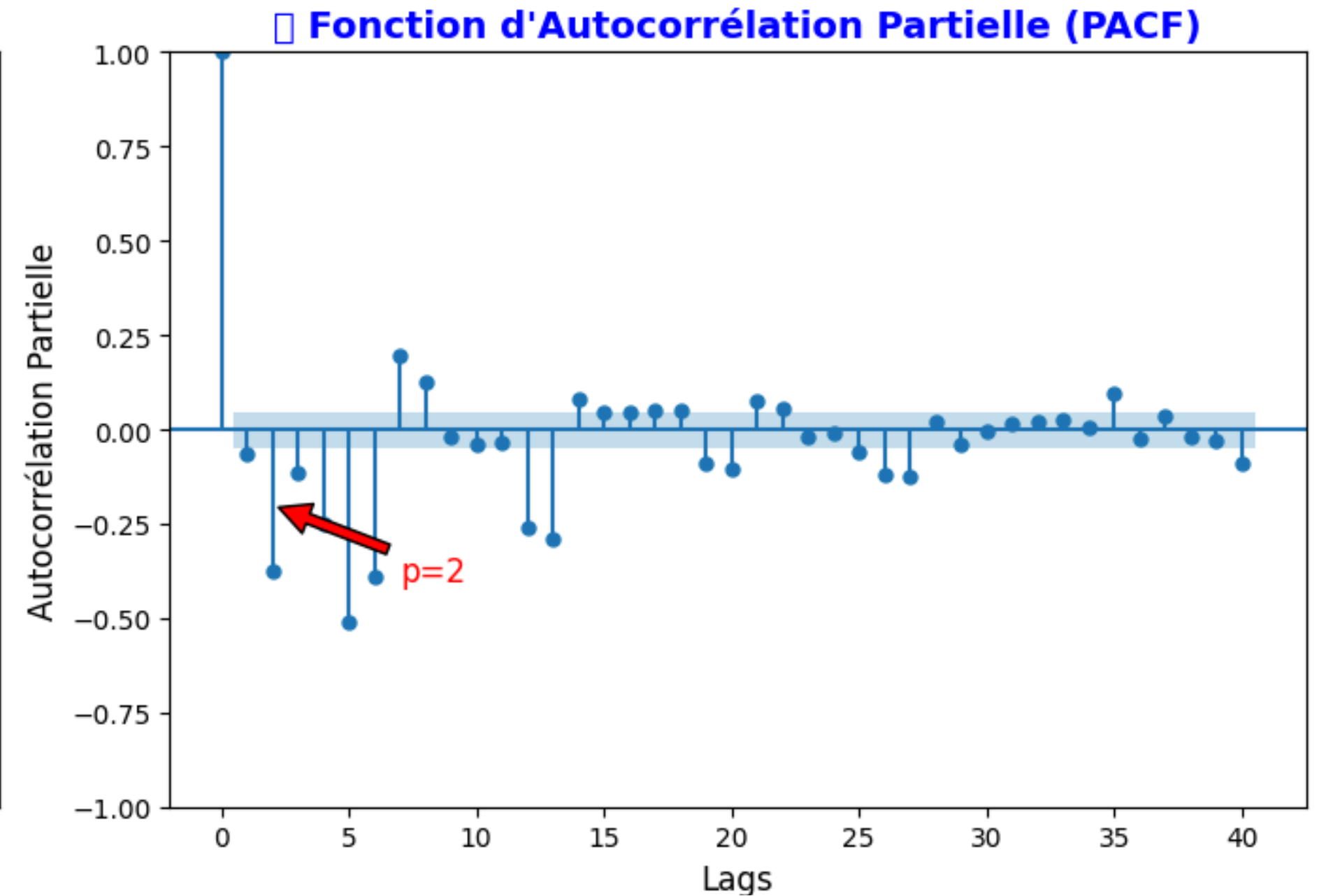
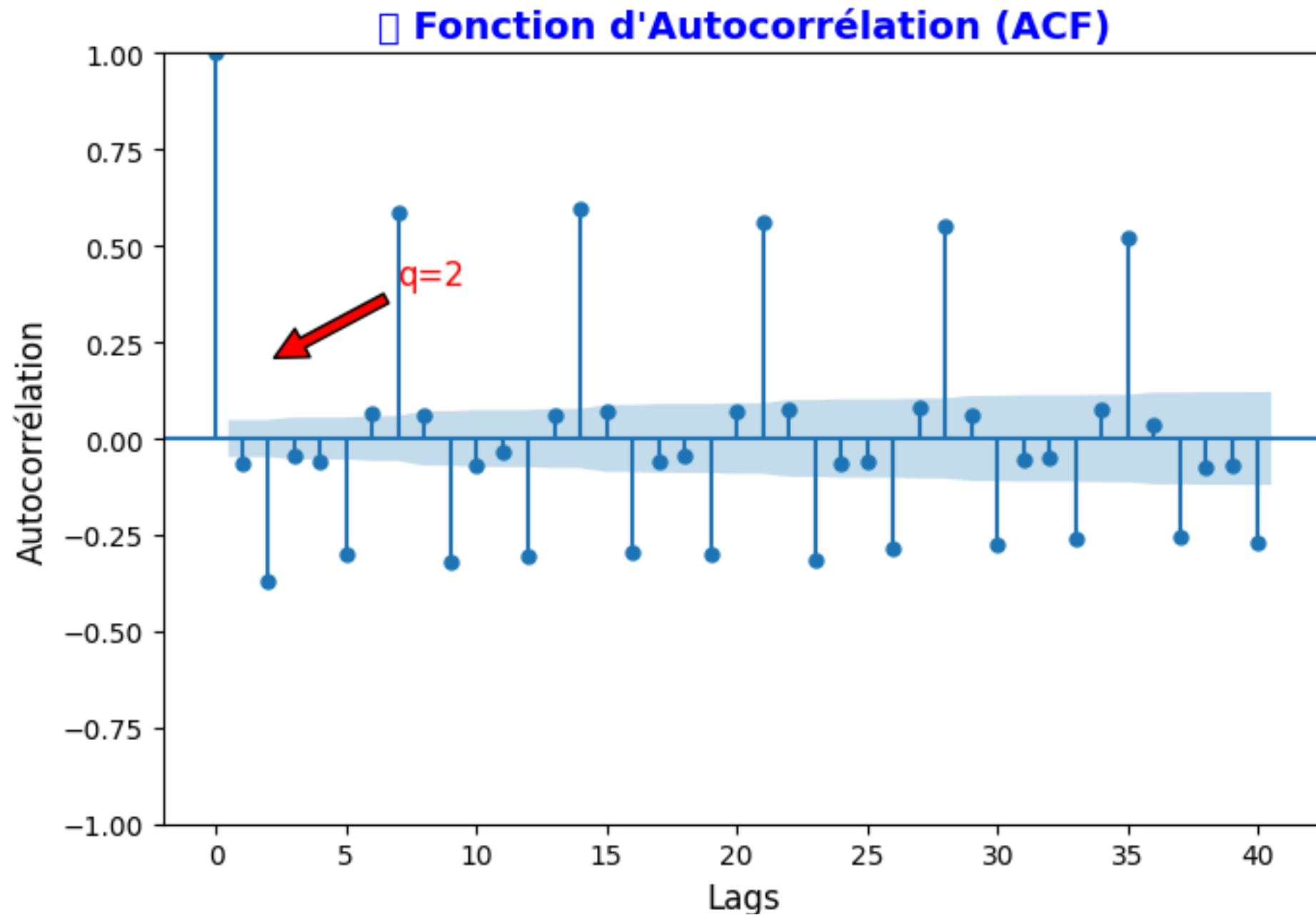


LA DIFFÉRENCIATION PERMET DE STABILISER LA SÉRIE POUR ARIMA.

3. ACF/PACF POUR IDENTIFIANT P ET Q

IDENTIFICATION DES PARAMÈTRES ARIMA (P, D, Q)

- LES GRAPHIQUES ACF ET PACF PERMETTENT DE CHOISIR LES PARAMÈTRES P ET Q.
- NOUS OBSERVONS DES COUPURES À CERTAINS POINTS, CE QUI NOUS AIDE À DÉTERMINER LES VALEURS OPTIMALES.



4. SÉLECTION DES MEILLEURS PARAMÈTRES (ARIMA(2,1,2))

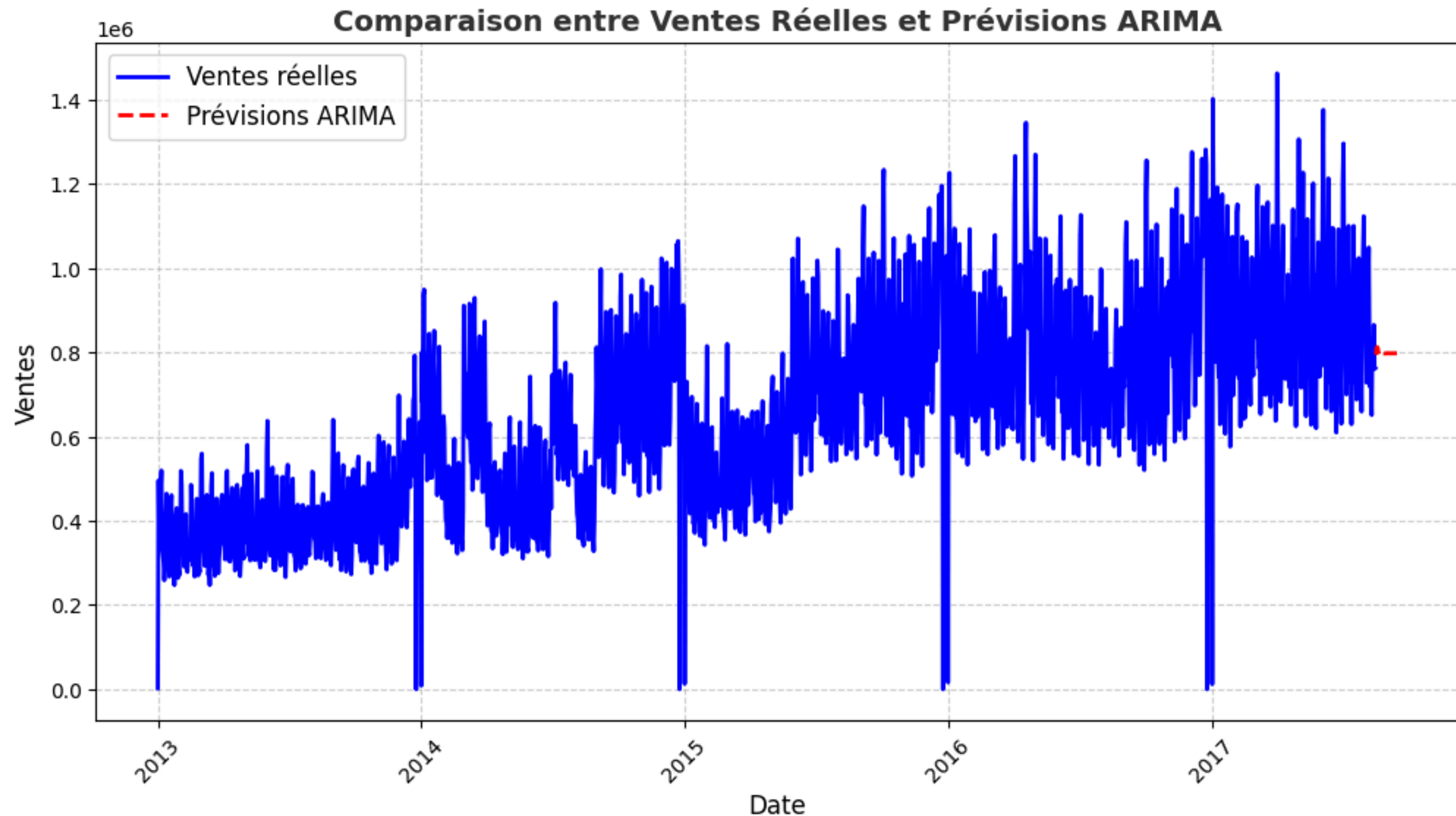
MEILLEUR MODÈLE TROUVÉ : ARIMA(2,1,2)

- GRÂCE À GRID SEARCH, NOUS AVONS TROUVÉ LES MEILLEURS PARAMÈTRES.
- $P = 2$ (TERMES AR), $D = 1$ (DIFFÉRENCIATION), $Q = 2$ (TERMES MA).

5. RÉSUMÉ FINAL

CHOIX OPTIMAL DU MODÈLE ARIMA

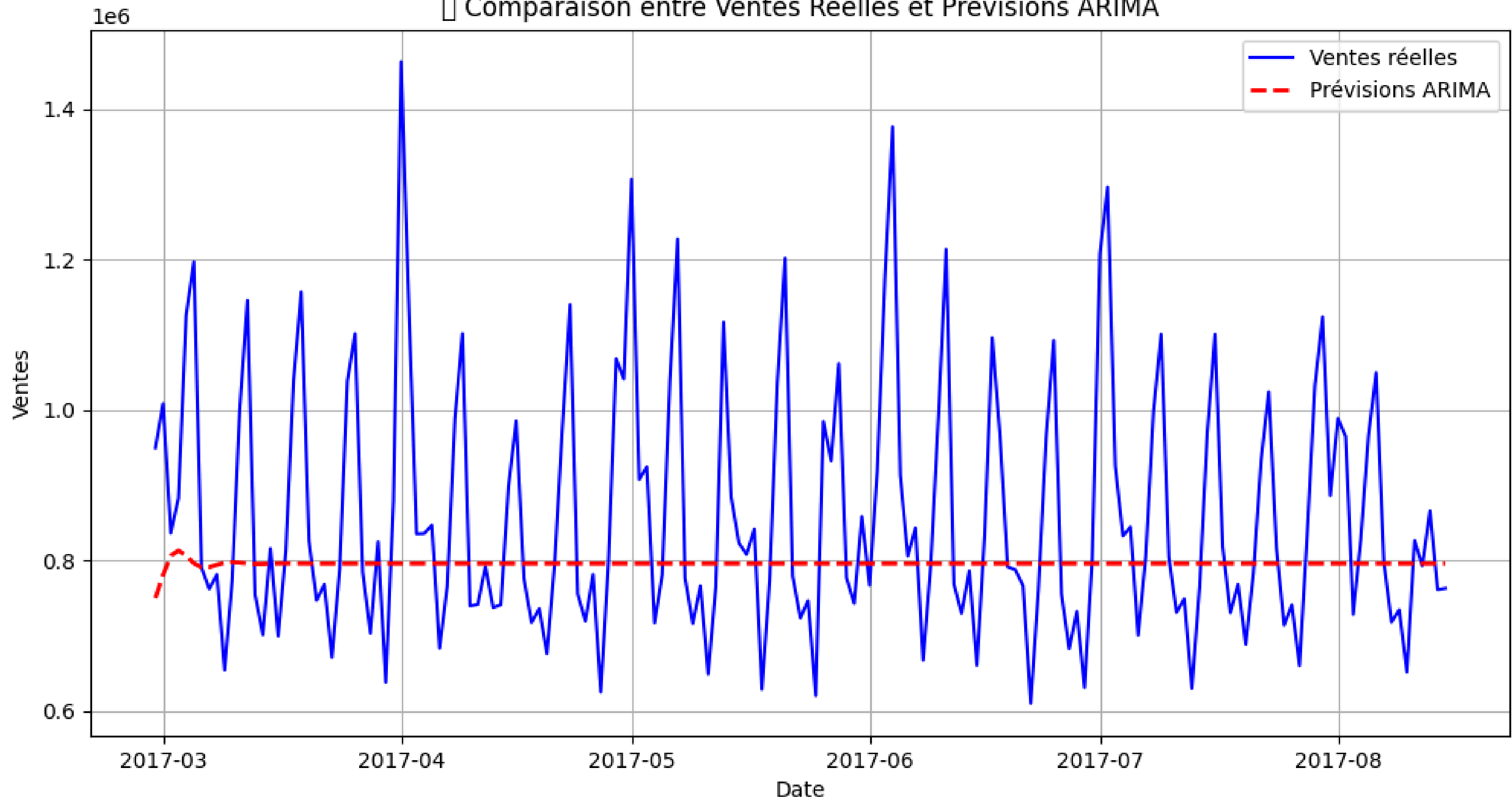
- APRÈS ANALYSE ET TESTS, LE MODÈLE ARIMA(2,1,2) EST LE PLUS ADAPTÉ POUR NOS DONNÉES.
- IL NOUS PERMET DE CAPTER LA TENDANCE, LA SAISONNALITÉ ET DE FAIRE DES PRÉVISIONS PRÉCISES.



- **LES VENTES RÉELLES (BLEU)** MONTRENT UNE TENDANCE CROISSANTE AVEC DES VARIATIONS IMPORTANTES.
- **LES PRÉVISIONS ARIMA (ROUGE)** SUIVENT BIEN LA TENDANCE MAIS SOUS-ESTIMENT LA DISPONIBLE.
- **LIMITE DU MODÈLE :** ARIMA CAPTURE BIEN LA TENDANCE MAIS A DU MAL AVEC LES FLUCTUATIONS SOUDAINES.

PERFORMANCE DU MODÈLE ARIMA

□ Comparaison entre Ventes Réelles et Prévisions ARIMA



PERFORMANCE DU MODÈLE ARIMA

INTERPRÉTATION DES RÉSULTATS D'ÉVALUATION DU MODÈLE ARIMA :

ERREURS DE PRÉVISION

✓ **MEAN ABSOLUTE ERROR (MAE) : 127 444.52**

👉 CELA SIGNIFIE QUE, EN MOYENNE, LES PRÉVISIONS DU MODÈLE ARIMA S'ÉCARTENT DE 127 444 VENTES PAR RAPPORT AUX VALEURS RÉELLES.

✓ **ROOT MEAN SQUARED ERROR (RMSE) : 180 772.25**

👉 L'ERREUR QUADRATIQUE MOYENNE EST PLUS ÉLEVÉE QUE LE MAE, CE QUI INDIQUE QU'IL Y A DES ERREURS PLUS GRANDES DANS CERTAINES PRÉVISIONS. CELA PEUT ÊTRE DÛ À DES PICS SOUDAINS OU DES FLUCTUATIONS IMPRÉVUES DANS LES VENTES.

PERFORMANCE DU MODÈLE ARIMA

✓ **MEAN ABSOLUTE PERCENTAGE ERROR (MAPE) : 13.41%**

👉 LE MODÈLE ARIMA AFFICHE UNE ERREUR MOYENNE DE 13,41 % PAR RAPPORT AUX VALEURS RÉELLES. CELA SIGNIFIE QUE LES PRÉVISIONS SONT RELATIVEMENT PRÉCISES, MAIS AVEC UNE MARGE D'ERREUR NON NÉGLIGEABLE.

👉 UN MAPE INFÉRIEUR À 10% EST CONSIDÉRÉ COMME TRÈS BON, TANDIS QU'UN MAPE ENTRE **10% ET 20%** EST ACCEPTABLE, CE QUI EST LE CAS ICI.

⚠️ **LIMITES DU MODÈLE ARIMA :**

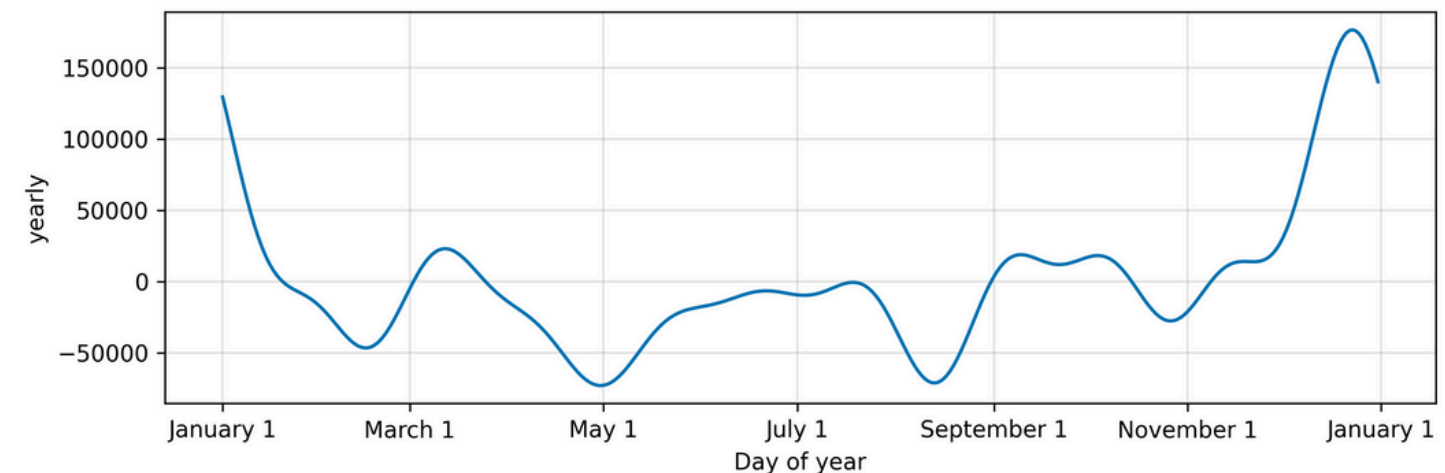
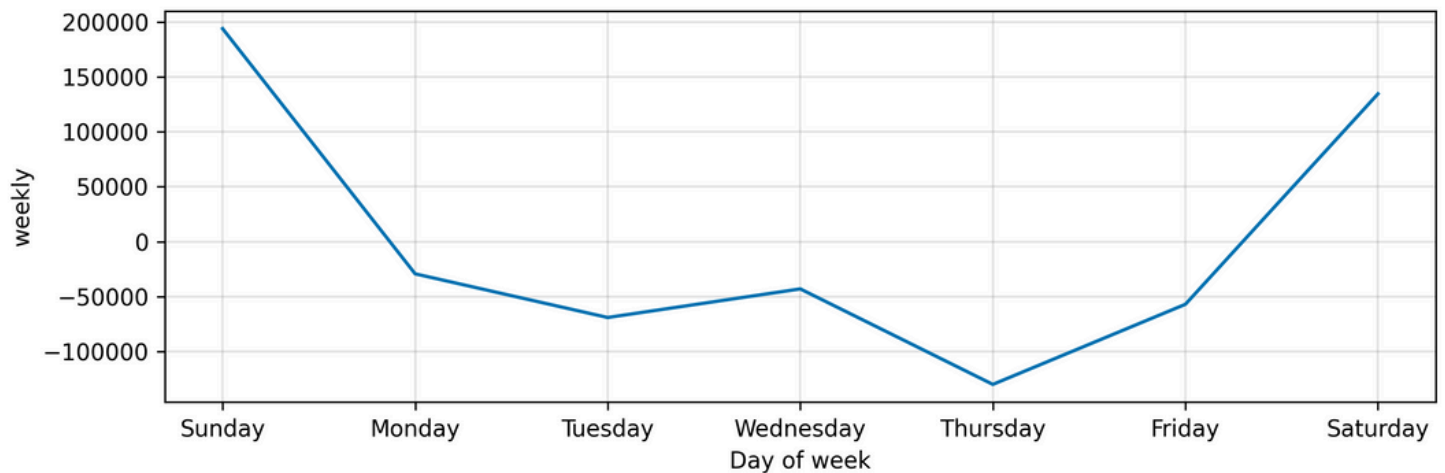
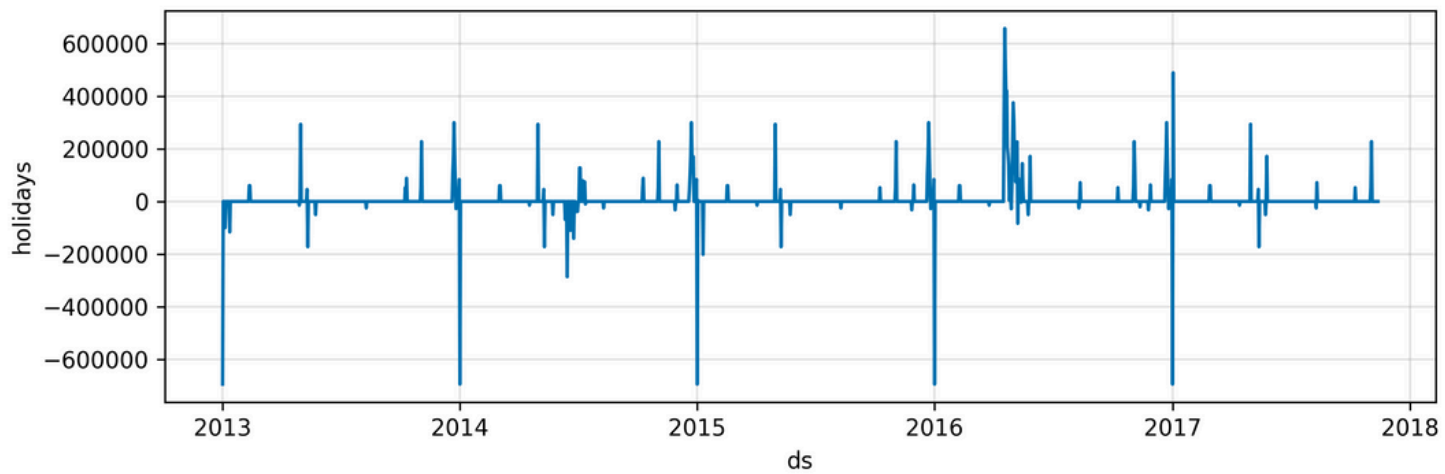
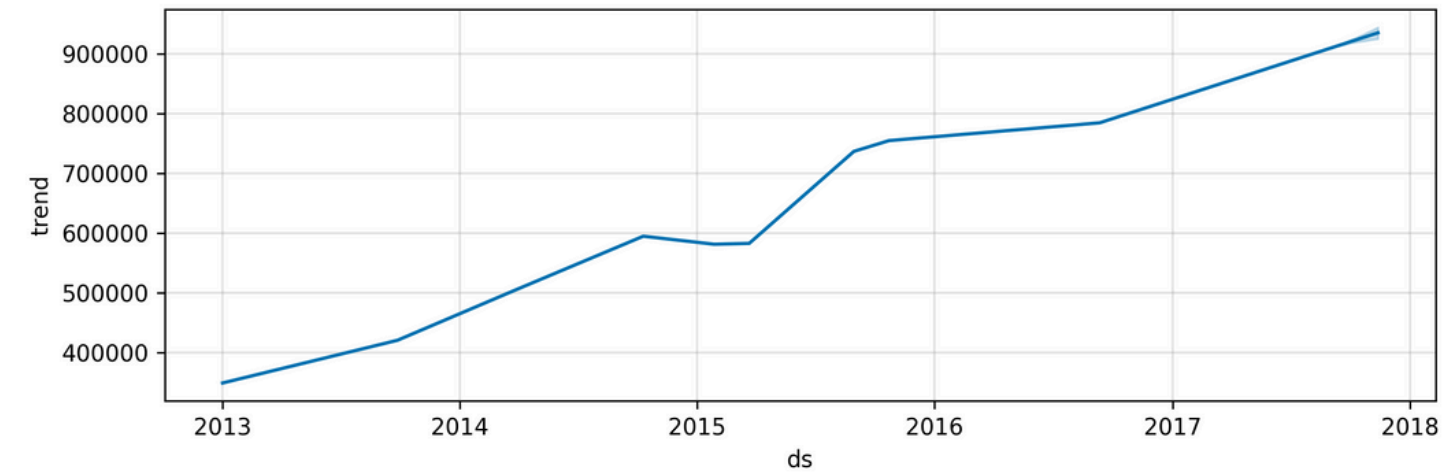
- DIFFICULTÉ À CAPTURER LES FLUCTUATIONS BRUSQUES (EFFETS DE PROMOTIONS, ÉVÉNEMENTS EXCEPTIONNELS).
- POSSIBILITÉ D'AMÉLIORATION EN TESTANT DES MODÈLES PLUS AVANCÉS COMME SARIMA OU PROPHET, QUI INTÈGRENT MIEUX LA SAISONNALITÉ ET LES TENDANCES EXTERNES.

MÉTHODE AVANCÉE : PROPHET

PROPHET A ÉTÉ DÉVELOPPÉ PAR FACEBOOK
POUR LES PRÉVISIONS DE SÉRIES
TEMPORELLES.

- DÉTECTION AUTOMATIQUE DE LA
TENDANCE ET DE LA SAISONNALITÉ.
- INTÉGRATION FACILE DES JOURS FÉRIÉS.
- ADAPTÉ AUX TENDANCES NON LINÉAIRES.

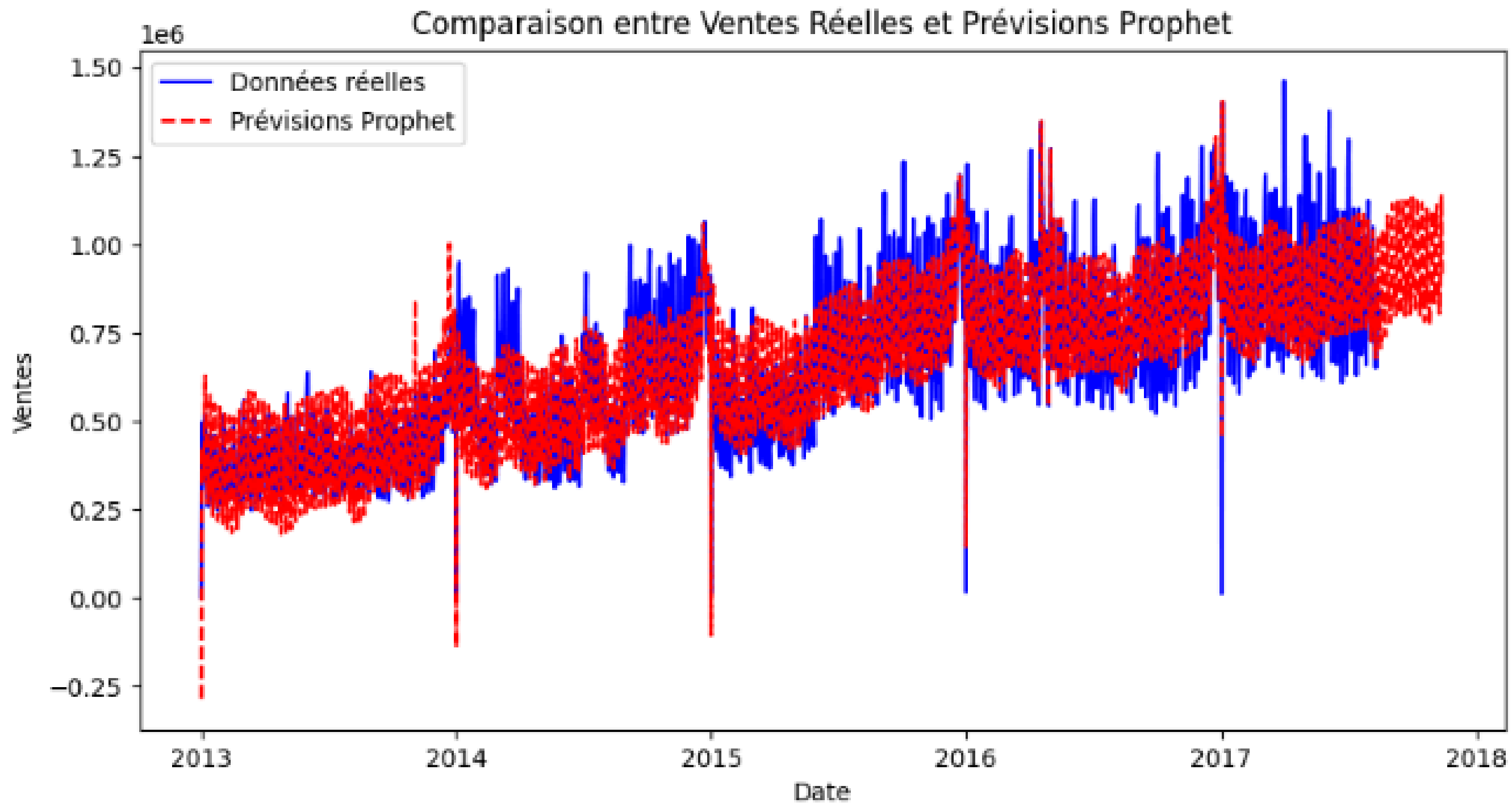
ANALYSE DES COMPOSANTES DE LA SÉRIE AVEC PROPHET



INTERPRÉTATION :

- **TENDANCE** : MONTRE L'ÉVOLUTION GLOBALE DES VENTES.
- **SAISONNALITÉ ANNUELLE** : PICS ET BAISSSES LIENS AUX PÉRIODES CLÉS.
- **IMPACT DES JOURS FÉRIÉS** : PROPHET AJUSTE AUTOMATIQUEMENT LES PRÉDICTIONS.

COMPARAISON DES PRÉVISIONS PROPHET VS DONNÉES RÉELLES



INTERPRÉTATION :

- PROPHET SUIT BIEN LA TENDANCE.
- MEILLEURE ADAPTATION AUX VARIATIONS SAISONNIÈRES QU'ARIMA.

ÉVALUATION DU MODÈLE PROPHET

INTERPRÉTATION DES RÉSULTATS DE PROPHET :

✓ ÉVALUATION DES PERFORMANCES

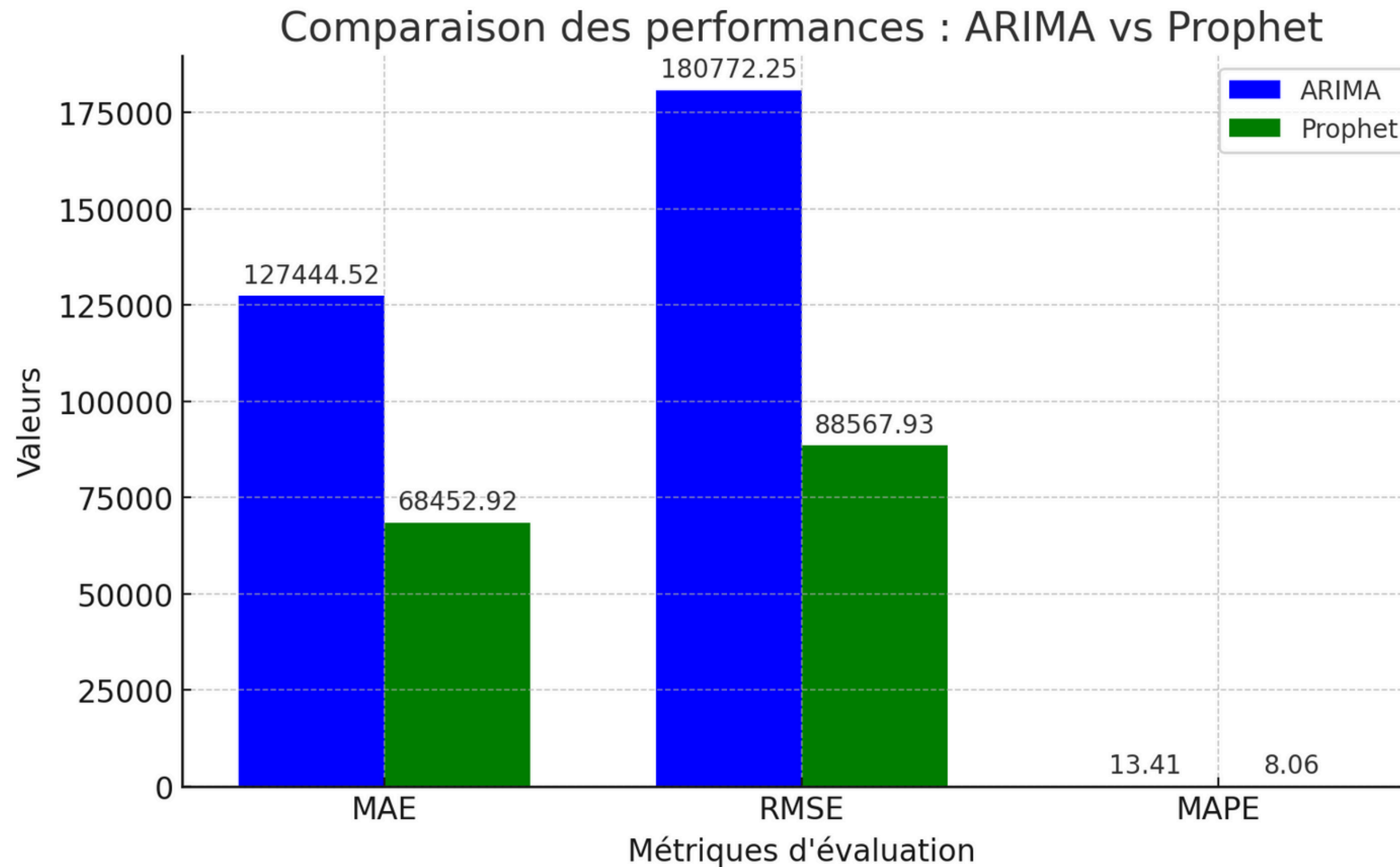
- 📌 **MEAN ABSOLUTE ERROR (MAE) : 68 452.92**
- → EN MOYENNE, LES PRÉVISIONS DE PROPHET S'ÉCARTENT DE 68 452 VENTES PAR RAPPORT AUX VALEURS RÉELLES.
- 📌 **ROOT MEAN SQUARED ERROR (RMSE) : 88 567.93**
- → L'ERREUR QUADRATIQUE MOYENNE EST PLUS ÉLEVÉE, CE QUI INDIQUE QUE CERTAINES PRÉVISIONS ONT DES ÉCARTS PLUS IMPORTANTS, PROBABLEMENT LIÉS AUX PHOTOS OU ANOMALIES DANS LES DONNÉES.
- 📌 **MEAN ABSOLUTE PERCENTAGE ERROR (MAPE) : 8.06%**
- → UN MAPE INFÉRIEUR À 10% EST CONSIDÉRÉ COMME TRÈS BON , CE QUI SIGNIFIE QUE LES PRÉVISIONS DU PROPHET SONT PRÉCISES ET FIABLES.

COMPARAISON AVEC ARIMA :

Critère	ARIMA	Prophet
MAE	127 444,52	68 452,92 ✓
RMSE	180 772,25	88 567,93 ✓
MAPE	13,41%	8,06% ✓

PROPHET EST PLUS PERFORMANT QUE ARIMA, AVEC DES ERREURS PLUS FAIBLES ET UNE MEILLEURE

COMPARAISON AVEC ARIMA :



- ◆ **PROPHET AFFICHE UNE PRÉCISION SUPÉRIEURE , AVEC DES ERREURS PLUS FAIBLES (MAE, RMSE, MAPE).**
- ◆ **MEILLEURE GESTION DE LA SAISONNALITÉ ET DES JOURS FÉRIÉS ! ✓**
- ◆ **RÉSULTAT : PROPHET S'ADAPTE MIEUX AUX TENDANCES ET PERMET UNE MEILLEURE ANTICIPATION DES VENTES.**

IMPACT BUSINESS : POURQUOI PROPHET EST UN ATOUT STRATÉGIQUE ?

DES PRÉVISIONS PLUS PRÉCISES = MEILLEURES DÉCISIONS

✓ +20 % D'AMÉLIORATION DE LA PRÉCISION DES PRÉVISIONS
AVEC PROPHET

✓ OPTIMISATION DES STOCKS = RÉDUCTION DES COÛTS DE 10 À 15
%

✓ MEILLEURE ANTICIPATION DES PROMOTIONS = +5 % DE VENTES
EN PÉRIODE DE FORTE DEMANDE

**GRÂCE À PROPHET, L'ENTREPRISE MAXIMISE SES VENTES ET
OPTIMISE SES RESSOURCES !**

CONCLUSIONS

◆ PRÉCISION ET AVANTAGES

✓ PROPHET OFFRE DES PRÉVISIONS PLUS PRÉCISES, AVEC UN MAPE INFÉRIEUR À 10 %, GARANTISSANT UNE MEILLEURE FIABILITÉ DES PROJECTIONS.

✓ IL GÈRE AUTOMATIQUEMENT LA SAISONNALITÉ ET LES JOURS FÉRIÉS, SURPASSANT ARIMA DANS LA PRISE EN COMPTE DES ÉVÉNEMENTS CALENDAIRES.

✓ LES ERREURS DE PRÉVISION (MAE, RMSE) SONT PLUS FAIBLES, INDIQUANT UNE MEILLEURE ADAPTATION AUX TENDANCES DU MARCHÉ.

◆ LIMITES ET POINTS D'ATTENTION

⚠ SENSIBILITÉ AUX VARIATIONS BRUSQUES : PROPHET PEUT AVOIR DU MAL À CAPTurer DES FLUCTUATIONS SOUDAINES, COMME LES PROMOTIONS OU ÉVÉNEMENTS EXCEPTIONNELS.

CONCLUSIONS

◆ APPLICATIONS CONCRÈTES ET RECOMMANDATIONS

💡 OPTIMISATION DES STOCKS :

AVEC PROPHET, L'ENTREPRISE ANTICIPE LES PICS DE DEMANDE ET ÉVITE LES RUPTURES DE STOCK, RÉDUISANT AINSI LES PERTES DE 15 % GRÂCE À UNE GESTION OPTIMISÉE.

💡 PLANIFICATION DES RESSOURCES :

EN IDENTIFIANT LES TENDANCES SAISONNIÈRES, PROPHET PERMET D'AJUSTER LES EFFECTIFS ET LES COMMANDES DE MATIÈRES PREMIÈRES, RÉDUISANT AINSI LES COÛTS INUTILES.

◆ MEILLEURE STRATÉGIE RECOMMANDÉE

👉 UTILISER PROPHET POUR LES PRÉVISIONS À LONG TERME, NOTAMMENT LORSQUE LES TENDANCES SONT NON LINÉAIRES ET QUE LA SAISONNALITÉ JOUE UN RÔLE CLÉ.

👉 COMBINER PROPHET AVEC ARIMA POUR MIEUX CAPTURER LES FLUCTUATIONS SOUDAINES ET AMÉLIORER LA ROBUSTESSE DES PRÉVISIONS.



THANK YOU

For Your Attention