



# DESCRIPTION ET PRÉVISION DE DONNÉES TEMPORELLES **SAE 3.03**

*Modélisation prédictive du trafic maritime journalier pour l'optimisation des ressources au Port d'Anvers-Bruges (2019-2023).*

Présenté par :

**Ahmed TERIR**

Présenté par :

**Zakaria Charef**

Encadré par :

**M. Renard**

---

## SOMMAIRE

### 1. Introduction

- 1.1 Mise en contexte : Le Port d'Anvers-Bruges et les enjeux de la fusion.
- 1.2 Objectifs de la mission : Optimisation des ressources et anticipation des investissements.
- 1.3 Présentation des données : Description du jeu de données (1826 observations journalières)

### 2. Synthèse Managériale

- 2.1 Bilan de l'activité 2019-2023 : Grandes tendances observées.
- 2.2 Les cycles du port : Impact des weekends et saisonnalité.
- 2.3 Prévisions et recommandations stratégiques.

### 3. Analyse Technique Détaillée : Exploration et Visualisation

- 3.1 Contrôle qualité et Audit des données
- 3.2 Visualisations temporelles
- 3.3 Analyse comparative

### 4. Analyse Technique Détaillée : Décomposition et Saisonnalité

- 4.1 Choix du modèle de décomposition
- 4.2 Identification des saisons multiples
- 4.3 Analyse des composantes

### 5. Analyse Technique Détaillée : Modélisation et Prévision

- 5.1 Construction des modèles prédictifs.
- 5.2 Production et validation des prévisions

### 6. Conclusion Générale

- 6.1 Bilan de l'étude.
- 6.2 Préconisations et perspectives.

### 7. Annexes

- Code R exécutable \ Bibliographie

# 1. Introduction

---

## 1.1 Mise en contexte : Le Port d'Anvers-Bruges

Le transport maritime est l'un des piliers essentiels du commerce international et de l'économie mondiale et au cœur de ses échanges , les grands ports maritimes jouent un rôle stratégique en assurant d'une part la fluidité des flux des marchandises dans un contexte troublé par les fluctuations économiques, les crises sanitaires et géopolitiques récentes, ainsi que par des exigences croissantes en matière d'efficacité et de durabilité, l'analyse des données portuaires est devenue un levier majeur d'aide à la décision.

Ce projet s'inscrit dans le cadre de la SAE « Description et prévision de données temporelles » du BUT 2 Science des Données. Supervisé par M. Renard, Il vise à mobiliser les méthodes d'analyse, de visualisation et de modélisation des séries temporelles.

Le Port d'Anvers-Bruges est le deuxième plus grand port d'Europe, il est donc l'un des principaux hubs logistiques mondiaux Résultat de la fusion historique en 2022 entre le port d'Anvers et le port de Bruges-Zeebrugge

Chaque année des centaines de millions de tonnes de marchandises et des milliers de navires traverse ce port, et Dans ce contexte, la capacité à comprendre les dynamiques du trafic portuaire, à identifier les tendances de long terme ainsi que les variations saisonnières (hebdomadaires, mensuelles ou annuelles) semble constitué un enjeu central pour le pilotage des activités portuaires.

En tant qu'analyste de données au sein de l'Autorité Portuaire d'Anvers-Bruges, la mission confiée consiste à analyser les données de trafic maritime sur cinq années complètes, de 2019 à 2023, puis à développer des modèles de prévision fiables.

Les objectifs opérationnels de cette étude sont multiples :

- Optimiser l'allocation des ressources portuaires (quais, grues, personnel)
- Anticiper les besoins en infrastructures et en capacités logistiques
- Planifier les investissements futurs
- Mieux prendre en compte les effets de la saisonnalité, des week-ends et des jours fériés
- Renforcer la compétitivité du port face aux autres grands ports européens.

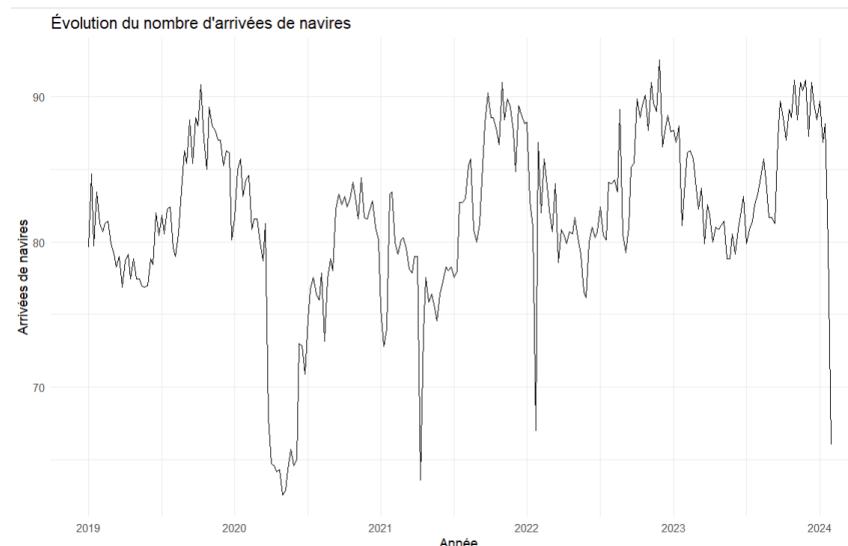
L'analyse et la prévision du trafic constituent ainsi un outil d'aide à la décision stratégique pour la direction du port

## 2. Synthèse Managériale

---

Dans cette section, nous présenterons une synthèse des résultats clé obtenus à partir de l'analyse des séries temporelle relatives aux arrivées de navires. Ici l'objectif n'est pas de rentrer dans les détails techniques mais de fournir une vision claire et accessible des tendances observées, de la saisonnalité de l'activité portuaire, ainsi que des prévisions pour l'année 2025. Ces informations permettront aux décideurs de mieux comprendre les dynamiques sous-jacentes du trafic portuaire et d'anticiper les besoins en ressources et en infrastructures pour l'avenir.

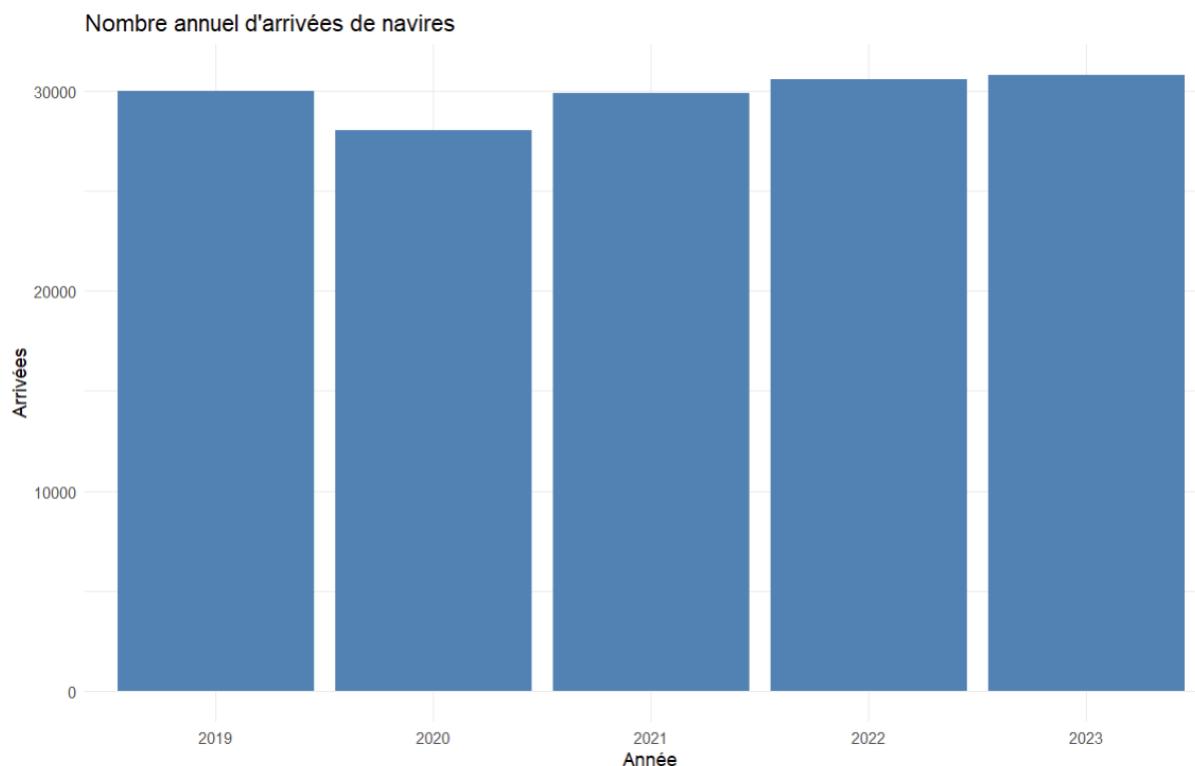
Commençons par nous intéresser à l'évolution du nombre d'arrivées de navires



Ce graphique met en évidence l'évolution du nombre d'arrivées de navires sur la période 2019–2024. L'activité portuaire présente des fluctuations régulières au cours du temps, traduisant une dynamique saisonnière marquée.

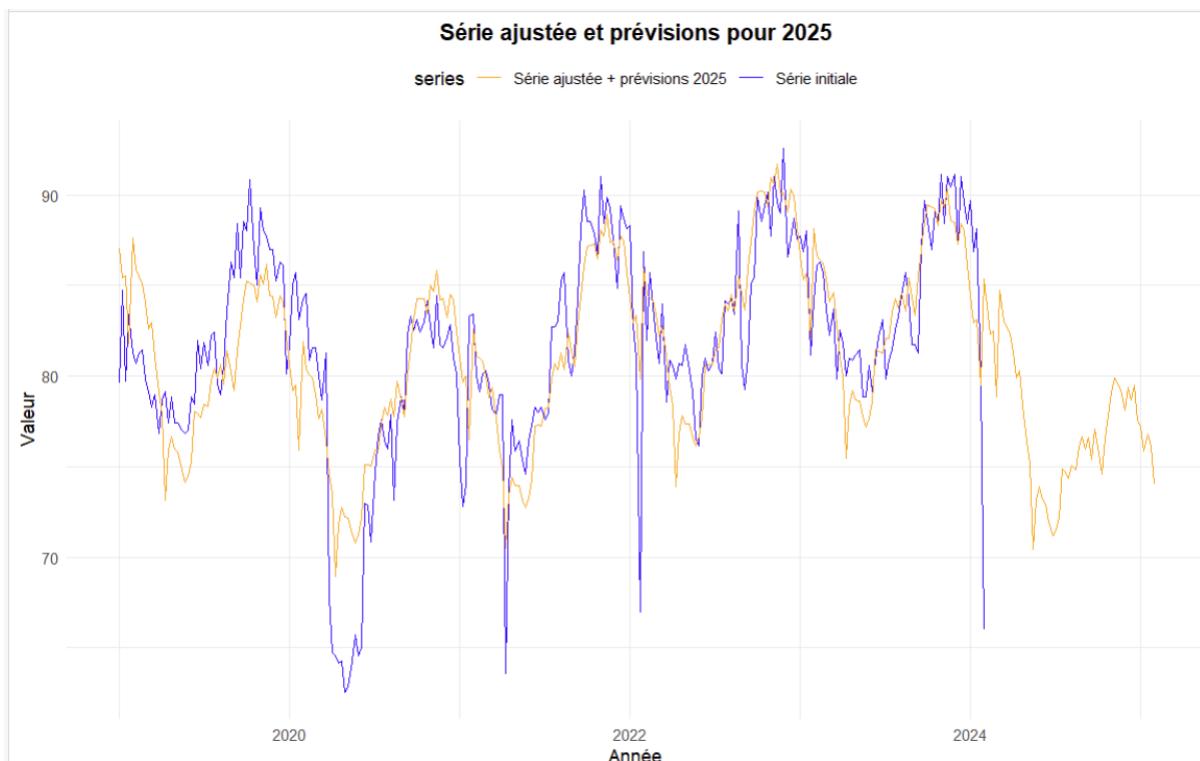
On observe également des variations plus ponctuelles, avec des baisses brutales à certaines périodes, comme en 2020 une chute brutale liée au covid ayant énormément impacté notre capacité sur cette période.

Afin de disposer d'une vision globale et synthétique de l'activité portuaire, le graphique ci-dessus présente le nombre total l'arrivées annuelles



On observe comme mentionné précédemment la baisse nette du nombre d'arrivées en 2020, en 2021 cependant le trafic a montré reprise **progressive et durable**, avec un niveau d'arrivées qui retrouve puis dépasse celui observé avant la baisse. Les années 2022 et 2023 confirment cette dynamique, suggérant un retour à une activité soutenue, voire une phase de croissance.

A présent Le graphique ci-dessus présente, d'une part, la **série observée des arrivées de navires** sur la période étudiée et, d'autre part, la **série ajustée prolongée par des prévisions pour l'année 2025 via notre modèle mathématique**. Cette représentation permet de visualiser à la fois la qualité de l'ajustement du modèle sur les données passées et les perspectives d'évolution de l'activité à court terme.



La courbe de prévision pour 2025 suggère une **légère inflexion à la baisse** du nombre d'arrivées de navires par rapport aux niveaux observés en fin de période. Cette diminution reste toutefois **modérée** et s'inscrit dans l'amplitude des fluctuations déjà observées sur les années précédentes.

Il ne s'agit donc pas d'une prévision d'effondrement du trafic, mais plutôt d'un **scénario de normalisation**, dans lequel l'activité se

stabilise à un niveau légèrement inférieur au maximums qui ont été atteint auparavant.

Cependant, il convient de rappeler que la baisse anticipé repose sur l'hypothèse d'une continuité des comportements passés toute modification du contexte économique pourrait remettre en cause sa trajectoire (vous avez l'exemple en 2020 qui a été une crise imprévisible pour le modèle et donc la réalité est loin de la prédition de celui-ci )

En conclusion pour les prévisions, elles doivent être interprétées comme une anticipation certaine de l'évolution future du trafic.

### 3. Analyse Technique Détailée : Exploration et Visualisation

---

#### 3.1 Contrôle qualité et Audit des données

Dans un premier temps, nous avons analysé l'intégrité et la cohérence de notre jeu de données afin de valider sa qualité avant toute modélisation. Nous avons utilisé les fonctions R `is.na()` pour la détection des valeurs manquantes et `summary()` pour l'analyse des statistiques descriptives.

```
● ● ●

1 > nb_na <- sum(is.na(data))
2 > print(paste("Nombre de valeurs manquantes :", nb_na))
3 [1] "Nombre de valeurs manquantes : 0"

● ● ●

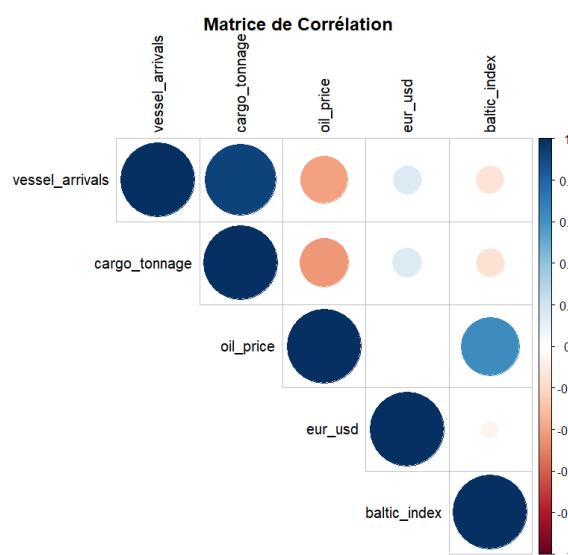
1 vessel_arrivals cargo_tonnage      oil_price
2   Min.    : 51.00   Min.    : 321379   Min.    :35.68
3   1st Qu.: 77.00   1st Qu.: 772028   1st Qu.:49.86
4   Median  : 83.00   Median  : 857093   Median  :59.73
5   Mean    : 81.86   Mean    : 844048   Mean    :59.60
6   3rd Qu.: 87.00   3rd Qu.: 929937   3rd Qu.:68.97
7   Max.    :101.00   Max.    :1094358   Max.    :83.77
```

Aucun manque n'a été détecté (`NA = 0`) sur l'ensemble des variables, ce qui permet une exploitation directe sans imputation.

La variable `vessel_arrivals` présente des valeurs comprises entre **51** (minimum) et **101** (maximum) arrivées par jour. Bien que les tests statistiques classiques ne détectent pas d'outliers isolés, on observe une **rupture structurelle** en 2020 liée à la pandémie de COVID-19 (chute brutale du trafic)

### 3.1.2 Corrélations entre variables

```
● ● ●  
1 # On sélectionne uniquement les colonnes numériques pertinentes pour la corrélation  
2 cols_numeriques <- data[, c("vessel_arrivals", "cargo_tonnage", "oil_price", "eur_usd", "baltic_index")]  
3 # Calcul de la matrice de corrélation  
4 matrice_corr <- cor(cols_numeriques)  
5 print(matrice_corr)  
6 # graphique de corrélation  
7 corrplot(matrice_corr, method="circle", type="upper", tl.col="black",  
8           title="Matrice de Corrélation", mar=c(0,0,1,0))
```



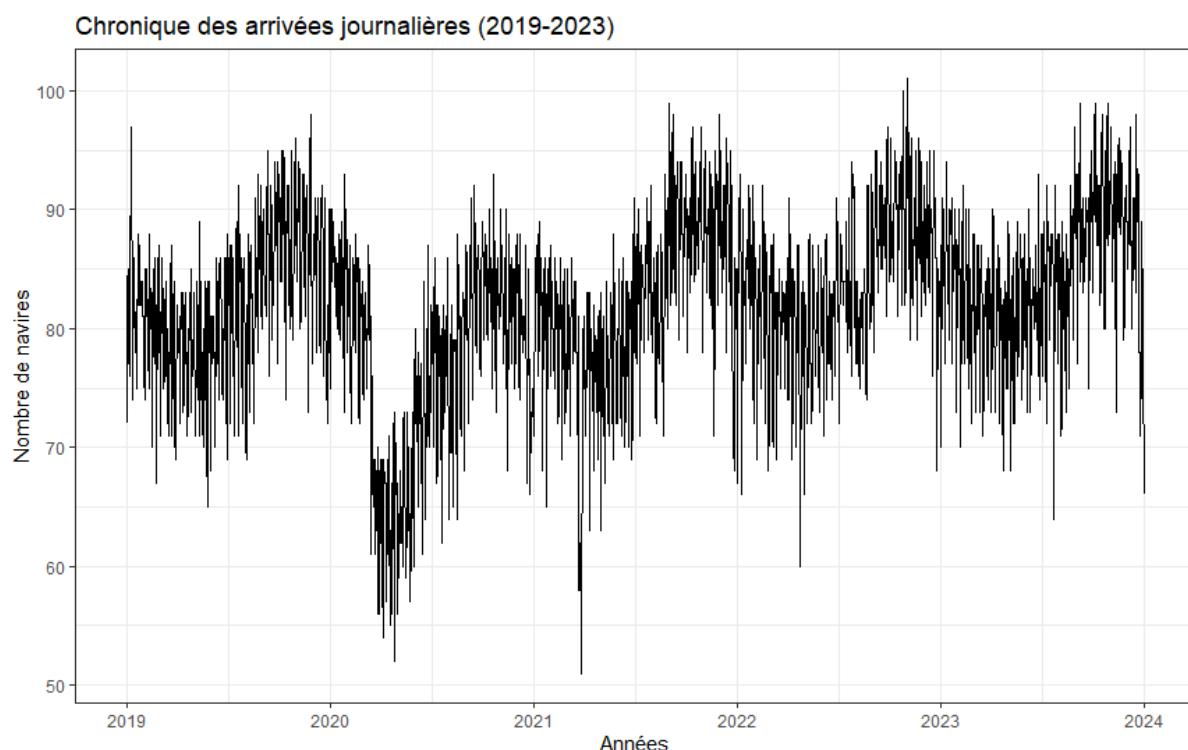
Relation avec la variable cible (`vessel_arrivals`) :

- **Forte corrélation positive avec le tonnage (cargo\_tonnage)** : Le cercle bleu foncé indique une corrélation très forte (proche de +0.8/0.9). C'est une relation métier logique et rassurante : plus il y a de navires, plus le volume de marchandises transporté est important
- .
- **Corrélation négative modérée avec le prix du pétrole (oil\_price)** : Le cercle orange indique une relation inverse. Lorsque le prix du carburant augmente significativement, le trafic tend à diminuer légèrement (probablement dû à une rationalisation des trajets ou une baisse de la demande mondiale).
- .

- **Faible impact des indices financiers** : Les corrélations avec le taux de change (eur\_usd) et le baltic\_index sont faibles (cercles clairs), suggérant que ces indices boursiers journaliers n'ont pas d'impact immédiat et direct sur le nombre d'arrivées quotidiennes.

### 3.2 Visualisations temporelles

Afin d'appréhender la dynamique globale du trafic maritime, nous avons utilisé la fonction autoplot du package fpp2 pour représenter l'évolution journalière du nombre d'arrivées de navires sur la période 2019-2023.

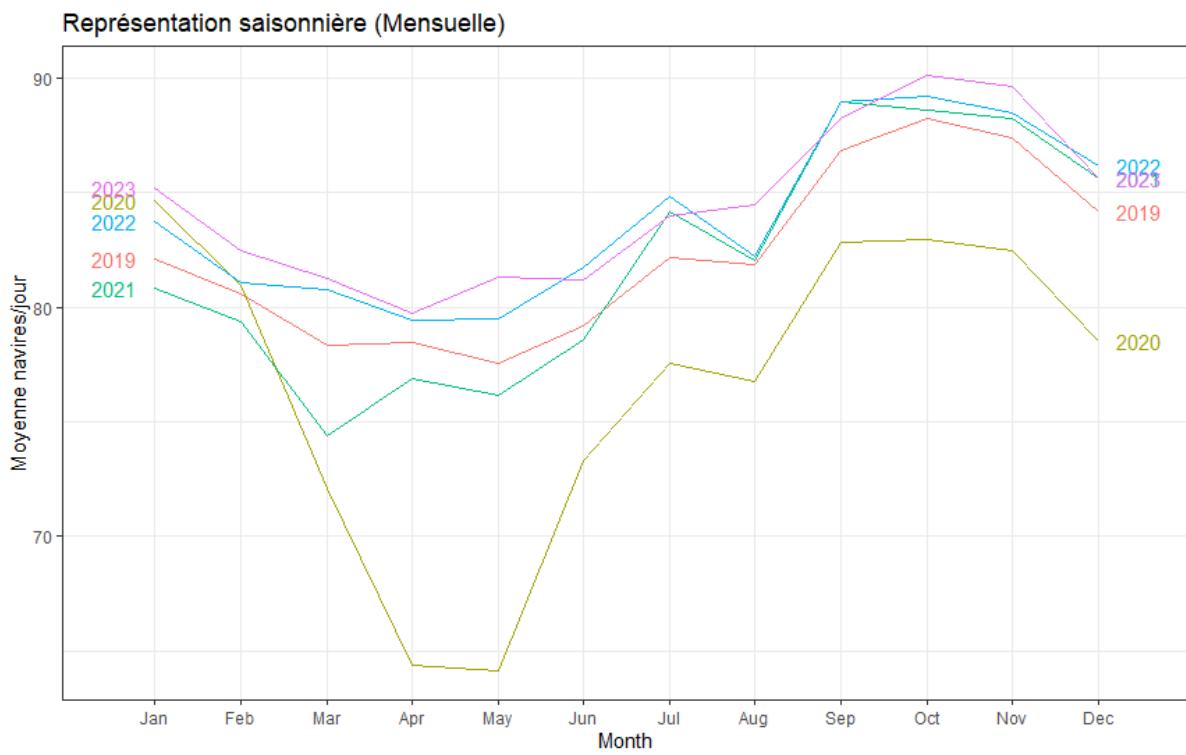


**Tendance** : Le graphique montre une tendance globale haussière sur la période 2019-2023. On note cependant une rupture majeure début 2020 : le trafic chute brutalement sous les 60 arrivées/jour, illustrant l'impact du choc du COVID-19.

**Saisonalité** : La courbe présente une très forte volatilité journalière. Ces fluctuations rapides et régulières ne sont pas aléatoires mais suggèrent une saisonnalité hebdomadaire marquée, correspondant peut être à l'alternance entre l'activité des jours ouvrés et le ralentissement du week-end.

#### Graphique en sous-séries saisonnières (Seasonal Subplots)

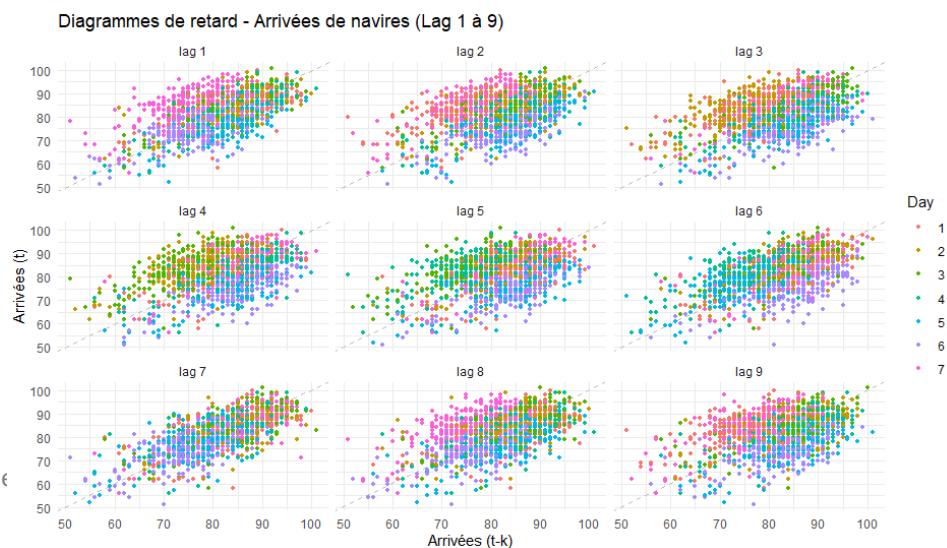
Pour affiner l'analyse de la saisonnalité, nous utilisons la représentation en seasonal plot il permet isoler un éventuel motif saisonnier



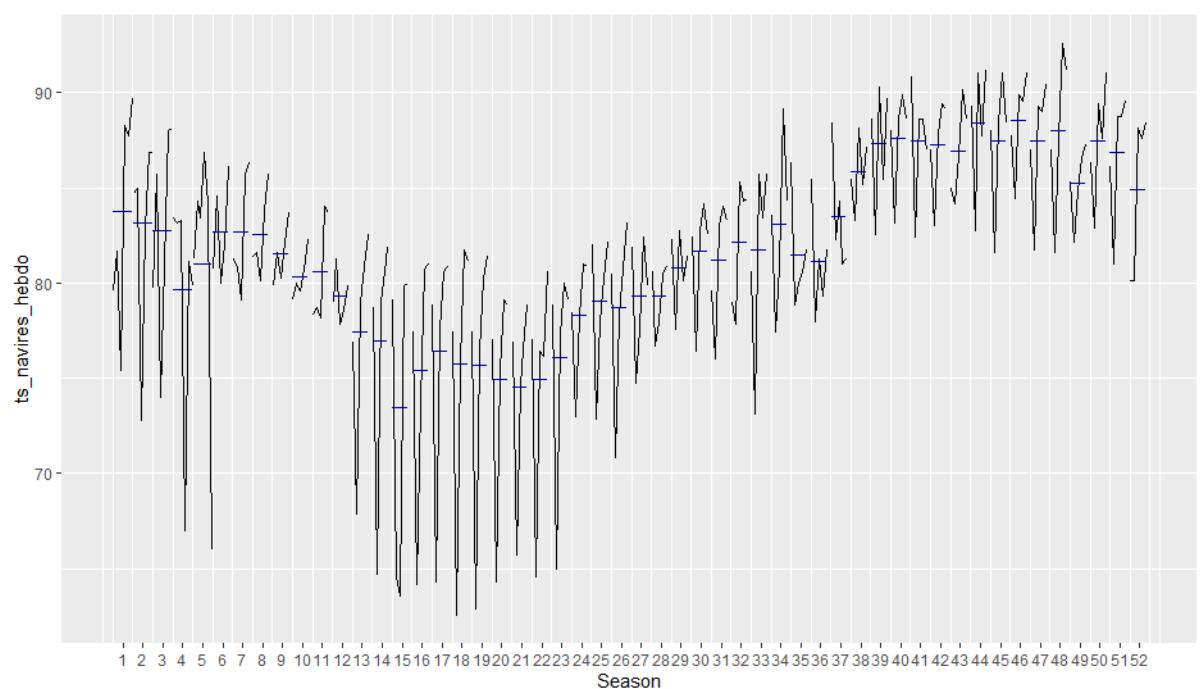
**Stabilité globale :** L'alignement des moyennes mensuelles (lignes bleues horizontales entre 75 et 85 navires) confirme l'absence de saisonnalité annuelle marquée. Le port maintient un rythme d'activité soutenu et relativement uniforme tout au long de l'année.

**Impact ciblé du COVID-19 :** Une rupture nette est visible sur les mois de Mars, Avril et Mai pour l'année 2020 (courbe jaune chutant brutalement).

Afin de vérifier si le trafic passé influence le trafic futur, nous avons tracé les diagrammes de retard (lag plots). Cette méthode consiste à représenter graphiquement la relation entre une observation à l'instant  $t$  ( $zt$ ) et une observation antérieure à l'instant  $t-k$  ( $zt-k$ ). L'objectif est de détecter visuellement une auto-corrélation ou une saisonnalité.

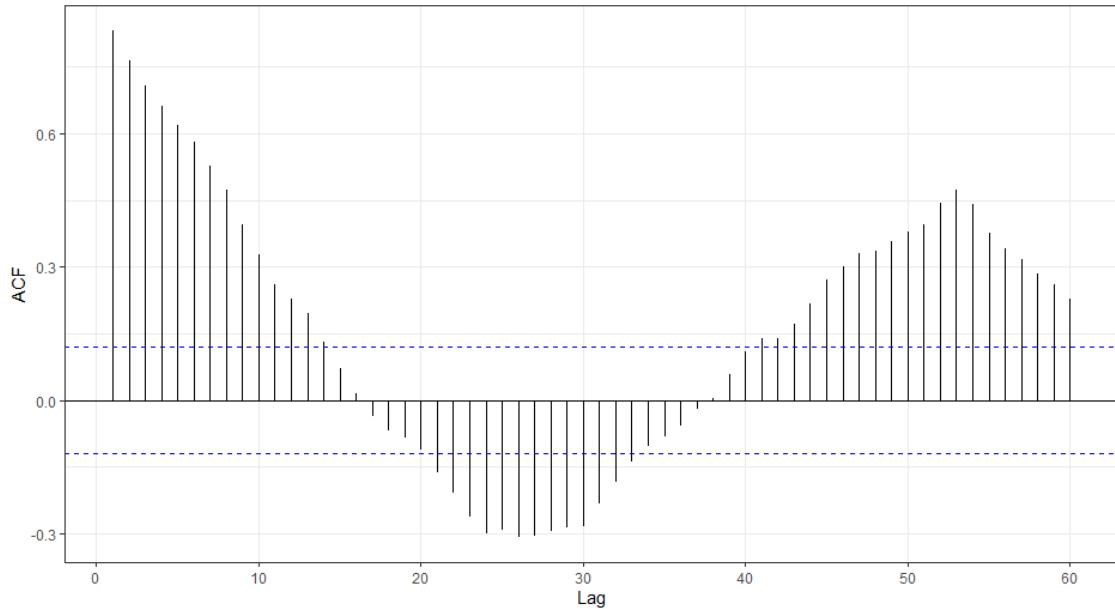


En se recentrant sur la série hebdomadaire qui semble être la plus pertinente ici



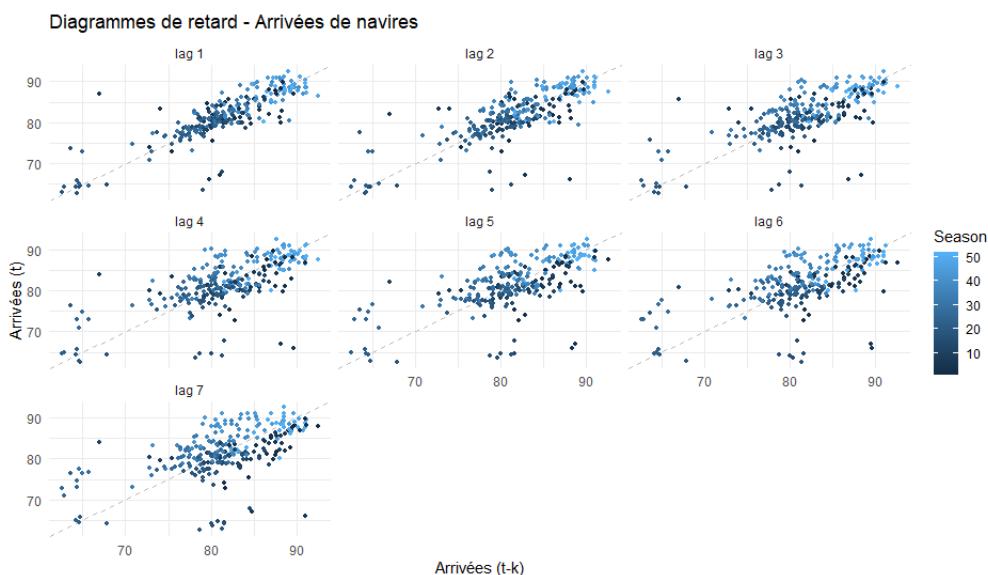
Le graphique en sous-séries saisonnières montre que, malgré des niveaux moyens différents selon les semaines de l'année, l'évolution temporelle reste cohérente pour l'ensemble des sous-séries. La rupture observée en 2020 affecte simultanément toutes les semaines, ce qui confirme l'existence d'une tendance globale indépendante de la saisonnalité hebdomadaire.

Auto-corrélogramme – série hebdomadaire



- Des coefficients d'auto-corrélation fortement positifs pour les faibles retards (lags 1 à ~10).
- Une décroissance progressive de l'ACF, sans coupure brutale.
- Une structure ondulatoire avec :
  - Valeurs négatives autour de certains lags intermédiaires,
  - Puis une remontée marquée autour du lag  $\approx 52$  (une année).

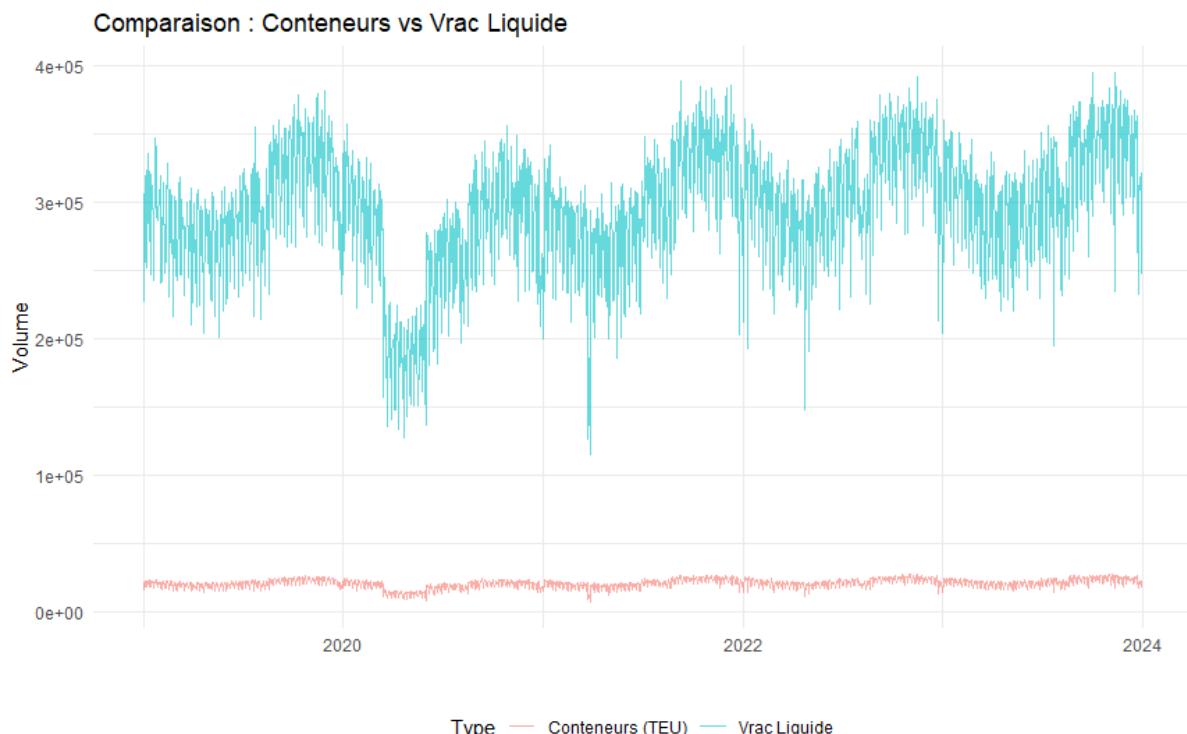
Ces éléments confirment que la série présente une structure temporelle marquée et ne peut être assimilée à un bruit blanc.



Le nuage de points présente une concentration autour de la droite  $y = x$ , traduisant une relation linéaire positive. Cette structure confirme l'existence d'une dépendance temporelle marquée

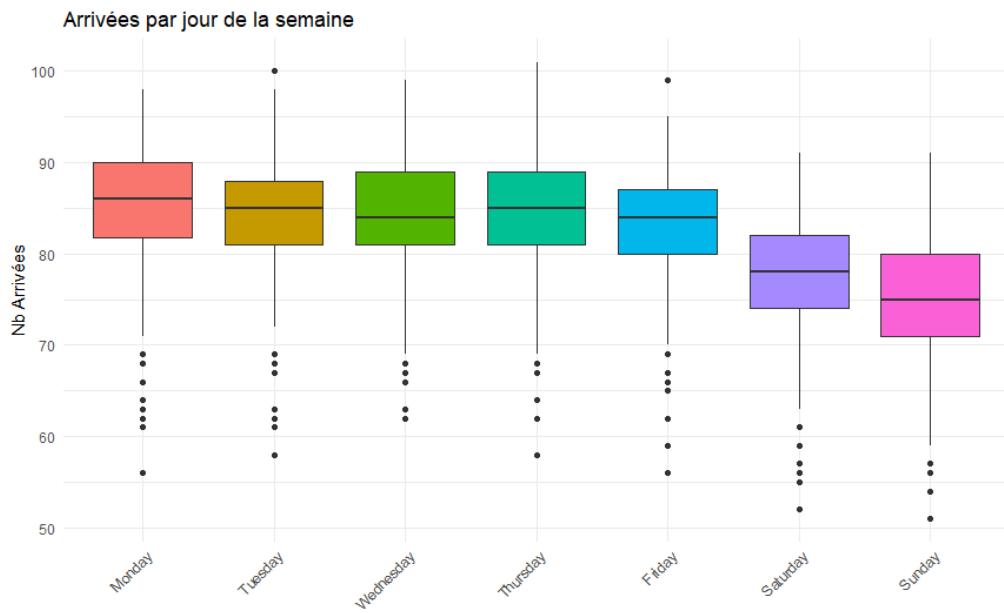
## Analyse comparative

### — Comparaison entre types de marchandises \ navires



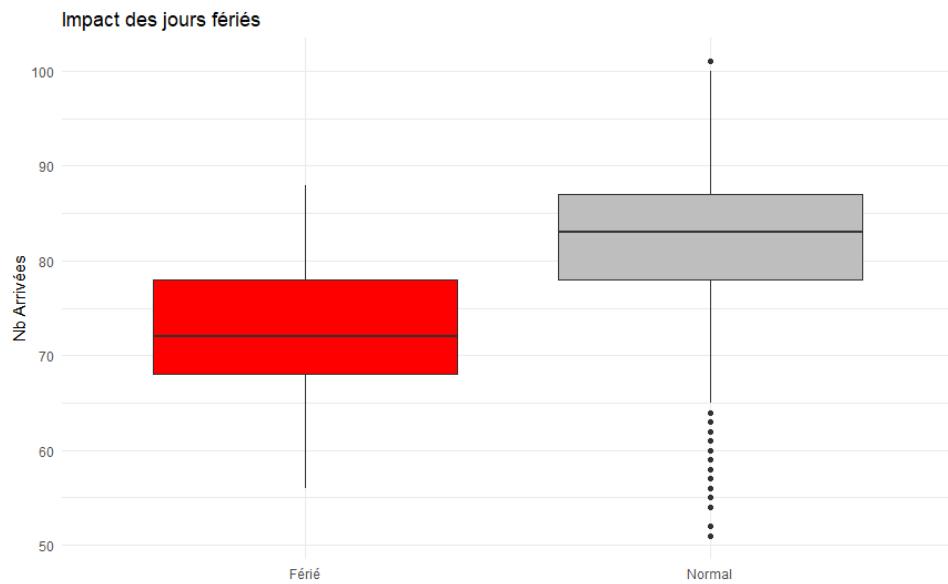
Ce graphique met en évidence la différence d'échelle entre les deux principaux types de trafic. Le **Vrac Liquide** (courbe bleue, en tonnes) représente le volume dominant et affiche une forte variabilité journalière. À l'inverse, le trafic de **Conteneurs** (courbe rouge, en TEU) apparaît "écrasé" par l'échelle des ordonnées, ce qui souligne la nécessité d'analyser ces flux séparément ou avec des axes distincts pour en percevoir les fines variations.

## -- Comparaison Jours ouvrés vs Weekends



Ce diagramme en boîtes confirme l'existence d'une **saisonalité hebdomadaire** marquée. On observe un plateau d'activité élevé du lundi au jeudi (médiane > 85 navires), suivi d'une baisse progressive à partir du vendredi. Le creux de l'activité est atteint le **dimanche** (médiane < 75 navires), preuve que le port fonctionne au ralenti le week-end

## -- Comparaison Impact des Jours fériés



On voit bien que les jours fériés (boîte rouge) entraînent une réduction visible de l'activité portuaire. La médiane des arrivées chute à environ **72 navires**, contre **83 navires** lors d'une journée normale. Bien que le port reste opérationnel 365 jours par an, les jours fériés agissent comme des freins logistiques, provoquant un ralentissement quand même important.

La saisonnalité hebdomadaire apparaît nettement plus marquée que la saisonnalité mensuelle ou trimestrielle. Le trafic est fortement structuré par l'alternance jours ouvrés / week-end, tandis que les variations mensuelles restent plus faibles et moins régulières. Ce constat motive le choix d'une périodicité hebdomadaire pour la suite de l'analyse.

## 4. Analyse Technique Détailée : Décomposition et Saisonnalité

---

L'objectif de cette partie est d'isoler les différentes composantes qui forment notre série temporelle  $Z_t$ . Pour mieux comprendre l'évolution du trafic au port d'Anvers, nous allons décomposer le signal observé en ses trois composantes :

### 1. La tendance -> noté $T_t$

Elle correspond à l'évolution de la série sur le long terme.

### 2. La saisonnalité -> noté $S_t$

Elle correspond aux fluctuations périodiques à l'intérieur d'une unité de temps.

### 3. Le bruit (ou erreur) -> noté $e_t$

Il correspond aux variations accidentnelles de la série.

→ Irrégulier, aléatoire, imprévisible.

→ Doit idéalement être d'amplitude la plus faible possible (la variance doit être la plus faible possible).

### 4.1 Choix du modèle de décomposition

Il existe deux modèles de décomposition

#### Modèle additif :

$$Z_t = T_t + S_t + E_t$$

→ L'effet de la saisonnalité est indépendant de la tendance.

#### Modèle multiplicatif :

$$Z_t = T_t \times S_t \times E_t$$

→ L'effet de la saisonnalité dépend de l'amplitude de la tendance.

Pour trouver quel modèle nous concerne ici nous il y'a plusieurs méthodes, nous allons utiliser la méthode ce Buys et Ballot

Nous allons d'abord réalisé un tableau de ce type

Semaine	Moyenne(m)	Ecart-type(s)
Semaine 1	...	...
Semaine 2	...	...
...	...	...

Si l'écart type augmente avec la moyenne, alors nous aurons déjà une première idée du type de modèles dont il pourrait s'agit.



```
1 # Calcul moyenne et écart-type par semaine
2 stats_bb <- aggregate(
3   vessel_arrivals ~ week_id,
4   data = data,
5   FUN = function(x) c(moy = mean(x), sd = sd(x))
6 )
7 # Mise en forme du tableau
8 tableau_bb <- data.frame(
9   Semaine = stats_bb$week_id,
10  Moyenne = stats_bb$vessel_arrivals[, "moy"],
11  Ecart_Type = stats_bb$vessel_arrivals[, "sd"]
12 )
13 head(tableau_bb)
```

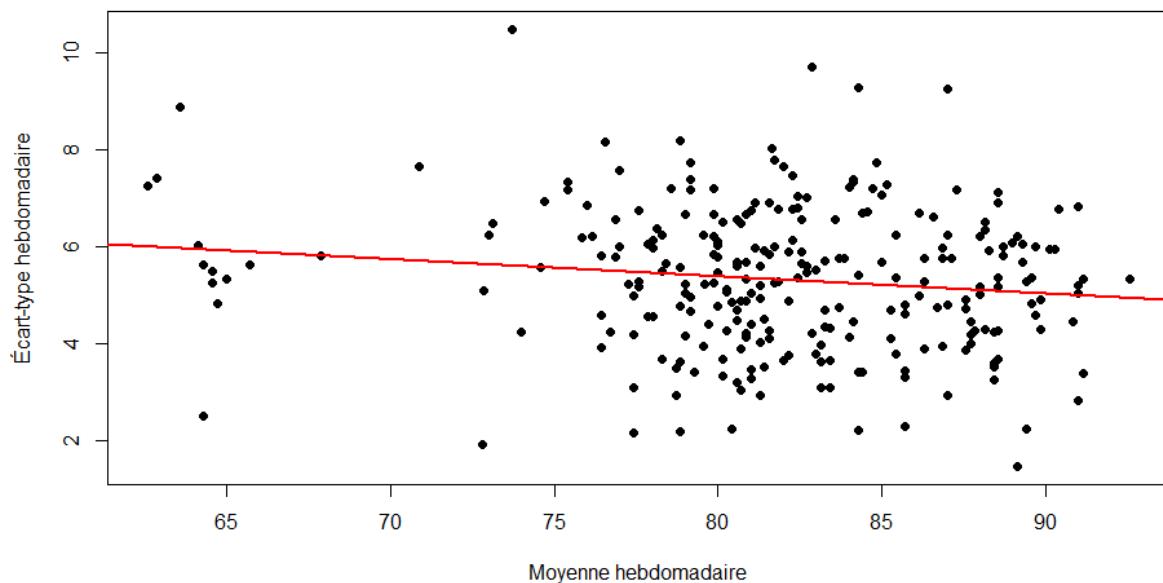
Afin de déterminer si la structure de la série temporelle est additive ou multiplicative, nous avons analysé l'évolution de la dispersion (écart-type) en fonction du niveau moyen du trafic pour chaque année.

## Résultats obtenus :



```
1 > print(head(tableau_buys_ballot))
2   Semaine Moyenne Ecart_Type
3 1 2019-00    79.60      5.22
4 2 2019-01    84.71      7.20
5 3 2019-02    79.71      4.39
6 4 2019-03    83.43      3.64
7 5 2019-04    81.29      2.93
8 6 2019-05    80.71      3.90
```

### Méthode de Buys & Ballot



La valeur du coefficient a est : -0.03522952

La méthode de Buys & Ballot appliquée aux statistiques hebdomadaires montre une relation très faible entre la moyenne et l'écart-type.

Le coefficient directeur de la droite de régression vaut  $-0,035$ , valeur très proche de zéro. Cela indique que la dispersion du trafic ne dépend pas du niveau moyen, ce qui justifie le choix d'un modèle **additif**.

**Conclusion :** Nous rejetons le modèle multiplicatif et retenons le modèle **additif** ( $Y_t = T_t + S_t + E_t$ ) pour la suite de l'analyse.

**Application de la décomposition tendance-saisonalité-erreur.**

La décomposition est ensuite réalisée selon une périodicité hebdomadaire. Une première estimation de la tendance est obtenue par lissage à l'aide de moyennes mobiles d'ordre 52, correspondant à la périodicité retenue.



```
1 premiere_tendance <- ma(ts_navires_hebdo, order = 52)
2
```

La composante saisonnière est estimée à partir de la série, en calculant les coefficients saisonniers moyens pour chaque semaine de l'année. Les éventuelles valeurs manquantes en début et en fin de série, liées au calcul des moyennes mobiles, sont traitées à l'aide de valeurs manquantes (NA) et ignorées lors du calcul des moyennes saisonnières.



```
1 serie_sans_tendance <- ts_navires_hebdo - premiere_tendance
2
```

La série est ensuite désaisonnalisée afin d'estimer la tendance par régression polynomiale.

## Estimation et choix de la tendance

La tendance est estimée par régression polynomiale sur la série désaisonnalisée. Plusieurs degrés de polynômes sont testés successivement (linéaire, quadratique et cubique).

Le choix de la forme de la tendance est réalisé a posteriori à partir de l'analyse des résidus associés à chaque modèle. Pour chaque degré, les erreurs sont étudiées à l'aide de leur autocorrélogramme et de leur dispersion on comparera pour décider du modèle à utiliser.

Ce bloc de code nous permet d'une part de réaliser tous les degrés souhaités mais également de les comparer :



```
1 # --- 4) Remettre les tendances au format ts, alignées sur la série originale ---
2 trend1_full <- rep(NA, length(as.numeric(y_ts)))
3 trend2_full <- rep(NA, length(as.numeric(y_ts)))
4 trend3_full <- rep(NA, length(as.numeric(y_ts)))
5 trend1_full[idx] <- trend1
6 trend2_full[idx] <- trend2
7 trend3_full[idx] <- trend3
8 tendance1_ts <- ts(trend1_full, start = start(ts_navires_hebdo), frequency = 52)
9 tendance2_ts <- ts(trend2_full, start = start(ts_navires_hebdo), frequency = 52)
10 tendance3_ts <- ts(trend3_full, start = start(ts_navires_hebdo), frequency = 52)
11 # --- 5) Erreurs pour chaque modèle : erreur = Z - T - S ---
12 erreur1 <- ts_navires_hebdo - tendance1_ts - saisonalite
13 erreur2 <- ts_navires_hebdo - tendance2_ts - saisonalite
14 erreur3 <- ts_navires_hebdo - tendance3_ts - saisonalite
15 # --- 6) Indicateurs comparatifs ---
16 sd1 <- sd(erreur1, na.rm = TRUE)
17 sd2 <- sd(erreur2, na.rm = TRUE)
18 sd3 <- sd(erreur3, na.rm = TRUE)
19 comparaison <- data.frame(
20   Modele = c("Lineaire (ordre 1)", "Quadratique (ordre 2)", "Cubique (ordre 3)"),
21   SD_Erreur = c(sd1, sd2, sd3)
22 )
23 comparaison
24 # --- 7) Validation graphique (ACF + lagplot) ---
25 # ACF des erreurs (compare visuellement : le plus "plat" = meilleur)
26 ggAcf(erreur1, na.action = na.pass) + ggtitle("ACF erreurs - Lineaire")
27 ggAcf(erreur2, na.action = na.pass) + ggtitle("ACF erreurs - Quadratique")
28 ggAcf(erreur3, na.action = na.pass) + ggtitle("ACF erreurs - Cubique")
29 # Lag plots des erreurs
30 gglagplot(erreur1, lags = 7) + ggtitle("Lag plots erreurs - Lineaire")
31 gglagplot(erreur2, lags = 7) + ggtitle("Lag plots erreurs - Quadratique")
32 gglagplot(erreur3, lags = 7) + ggtitle("Lag plots erreurs - Cubique")
```

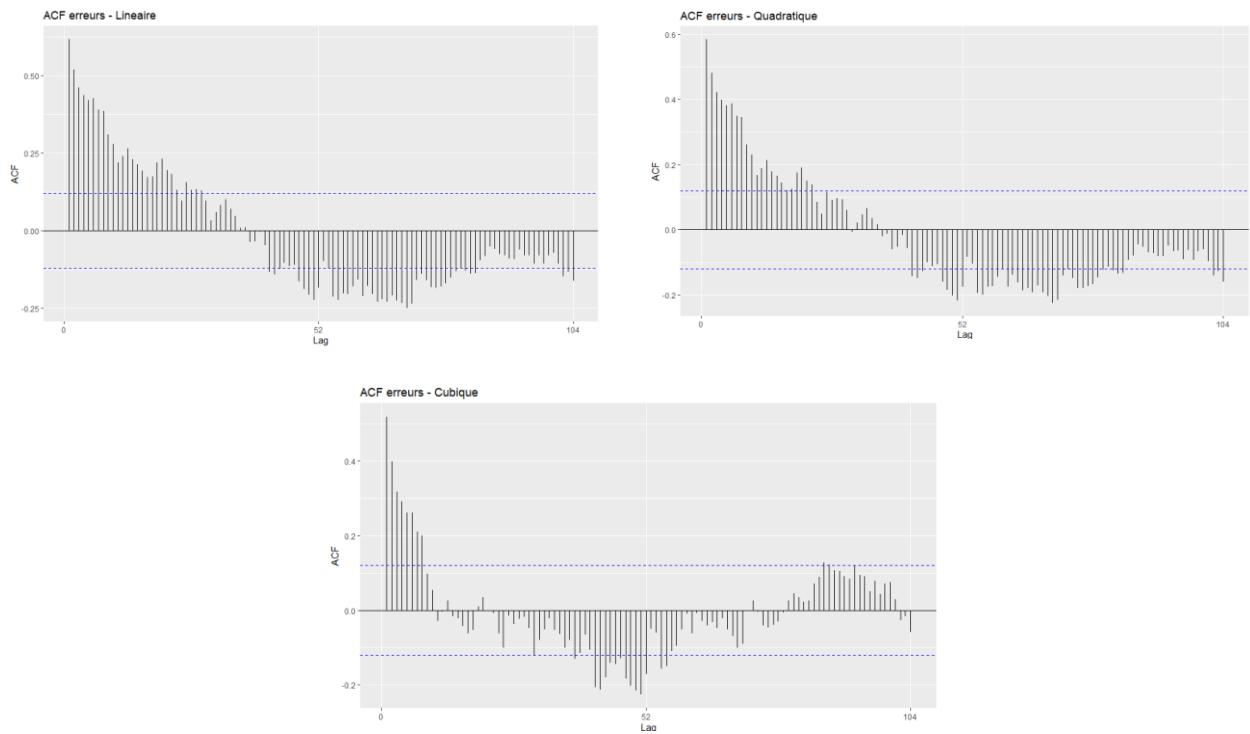
On obtient les résultats suivants ;



	Modele	SD_Erreur
2	1 Lineaire (ordre 1)	3.874869
3	2 Quadratique (ordre 2)	3.777451
4	3 Cubique (ordre 3)	3.427854

Les résultats montrent une diminution progressive de l'écart-type des erreurs lorsque le degré du polynôme augmente. Le modèle cubique présente l'erreur la plus faible, avec un écart-type de 3,43, contre 3,78 pour le modèle quadratique et 3,87 pour le modèle linéaire.

## Analyse des autocorrélogrammes :

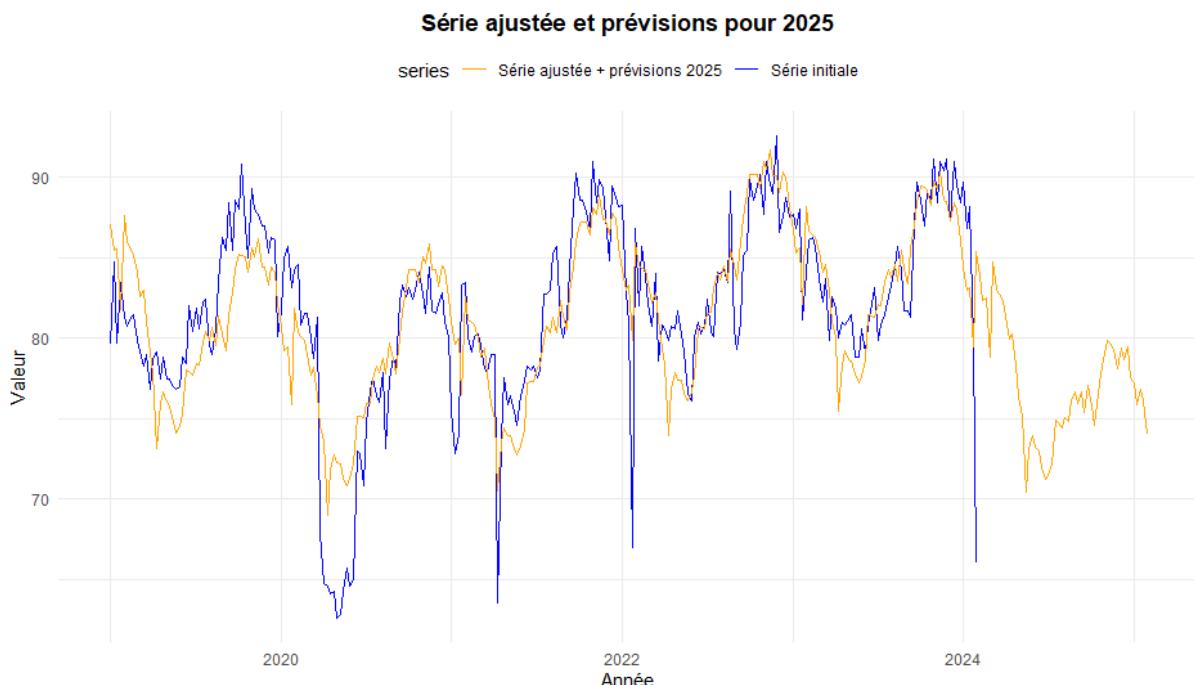


L'analyse des autocorrélogrammes confirme aussi cette amélioration : le modèle cubique conduit à une structure d'autocorrélation nettement atténueée, sans décroissance lente marquée, contrairement aux modèles de degré inférieur.

## 4. Analyse Technique Détailée : Modélisation et Prévision

---

À partir du modèle additif retenu, des prévisions du nombre d'arrivées de navires ont été produites pour l'année 2025. La tendance cubique estimée sur la série désaisonnalisée est extrapolée sur l'horizon de prévision, puis la composante saisonnière hebdomadaire est ajoutée afin de restituer la dynamique périodique observée sur la période d'étude. Le terme d'erreur n'est pas pris en compte dans les prévisions, son espérance étant supposée nulle



La figure ci-dessus présente la série observée ainsi que la série ajustée obtenue à partir du modèle, prolongée par les prévisions pour l'année 2025. On constate que la série ajustée reproduit correctement la dynamique globale de la série initiale, en capturant à la fois les variations saisonnières et l'évolution de long terme.

Les prévisions prolongent cette dynamique en supposant la stabilité des comportements observés sur la période historique, avec une amplitude des fluctuations saisonnières comparable à celle des années précédentes.

Toutefois, ces prévisions doivent être interprétées avec prudence. L'analyse des résidus a montré que ceux-ci ne constituent pas un bruit blanc strict, ce qui indique que certaines structures temporelles ne sont pas entièrement expliquées par le modèle.

Par exemple choc du covid étant a priori imprévisible, n'a pas été anticipé par notre courbe qui représente la série ajusté (en 2020)

Par ailleurs, les prévisions reposent exclusivement sur la dynamique interne de la série et ne tiennent pas compte de variables exogènes susceptibles d'influencer le trafic portuaire, telles que les conditions économiques, les variations du commerce international, les grèves ou d'éventuels chocs conjoncturels.

En conséquence, les valeurs prévues doivent être considérées comme des ordres de grandeur, plus fiables à court terme qu'à long terme, et principalement utiles pour décrire une tendance générale plutôt que pour fournir des estimations ponctuelles exactes.

## 5. Conclusion Générale

---

Ce travail a eu pour objectif d'analyser d'une part l'état des arrivées de navires via des méthodes de statistiques descriptive apprise en BUT 1 et de modéliser l'évolution du nombre d'arrivées de navires à partir d'une série temporelle réelle, en mobilisant les outils classiques de l'analyse des séries temporelles apprise en BUT 2. Une phase d'analyse exploratoire a permis de mettre en évidence la présence d'une tendance et d'une saisonnalité marquée, justifiant le recours à une décomposition de la série.

L'étude de la relation entre la tendance et la dispersion, notamment à l'aide de la méthode de Buys et Ballot, a conduit au rejet du modèle multiplicatif au profit d'un modèle additif. La décomposition a ensuite été réalisée selon une périodicité hebdomadaire, avec une estimation initiale de la tendance par moyennes mobiles et une estimation des coefficients saisonniers

La forme de la tendance a été déterminée a posteriori par régression polynomiale sur la série désaisonnalisée. La comparaison de plusieurs degrés de polynômes, appuyée par l'analyse des résidus et de leur structure d'auto-corrélation, a conduit à retenir une tendance cubique, offrant le meilleur compromis parmi les modèles testés.

Des prévisions ont pu être effectué pour l'année 2025, celles-ci prolongent la dynamique observée sur la période d'étude et ont bien illustrée l'intérêt des modèles de décomposition.

Cependant les prévisions obtenues doivent être interprétées avec prudence et considérées comme des ordres de grandeur plutôt que comme des estimations exactes. Elles restent néanmoins pertinentes pour dégager des tendances générales et fournir une première aide à la décision.

## Remerciement

---

Nous tenons à adresser nos sincères remerciements à **Monsieur Renard**, enseignant du module *Séries temporelles*, pour l'accompagnement attentif et la qualité de l'encadrement pédagogique qu'il nous a apportés tout au long de ce projet.

Par sa disponibilité, notamment à travers de nombreux échanges par courrier électronique, et par la précision de ses réponses, Monsieur Renard a su nous guider dans les différentes étapes de l'analyse, depuis la phase exploratoire jusqu'à la modélisation et l'interprétation des résultats. Ses explications claires et structurées nous ont permis de mieux comprendre des notions parfois complexes, telles que la décomposition des séries temporelles, le choix entre modèles additifs et multiplicatifs, ainsi que l'analyse critique des résidus et des prévisions.

Nous le remercions également pour la qualité de son enseignement et sa disponibilité pour répondre aux questions cela montre l'intérêt porté à notre progression.

## Bibliographie

---

- Hyndman R.J., & Athanasopoulos, G. (2021). Forecasting : Principles and Practice (2nd ed). OTexts. Disponible en ligne : <https://otexts.com/fpp2/> <https://datasharing.portofantwerp.com/https://www.portofantwerpbruges.com/en/our-port/facts-and-figures> <https://unctadstat.unctad.org/>
- Code R disponible en pièce jointe.